NASA CR-143681

OPERATIONS RESEARCH, Inc.

SILVER SPRING, MARYLAND

STUDY OF ADVANCED SYSTEM FOR POSITION LOCATION AND NAVIGATION FINAL REPORT

AUGUST 1972

Prepared under Contract NAS5-21685
for
National Aeronautics and Space Administration
Goddard Space Flight Center
Greenbelt, Maryland 20771

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

1. INTRODUCTION AND SUMMARY

Two measures of the relative performance of satellite based data collection and location systems are the precision with which platforms can be located (and/or their velocity estimated) and system capacity (average bit rate handled by the satellite). This report investigates two concepts whereby improvements in location precision and system capacity can be achieved relative to current systems. One of these concepts (TRILOC) can provide either or both improvements in location and system capacity by utilizing three measurements acquired during a single transmission from a platform for location purposes. The other concept is to increase system capacity by means of directive antennas on board the satellite for the purpose of reducing interference (either between platforms or from RFI) and/or reducing platform power.

The TRILOC concept is evaluated in Section 2. The location algorithm is developed assuming range-rate, radial-acceleration, and range differences are measured on-board the satellite during a transmission from a platform. An error analysis of this algorithm is then derived from which platform location errors are related to the precision with which the three measurements are made. An example TRILOC processor is then postulated to estimate the three measurement precisions.

Particular conclusions related to the TRILOC concept can be summarized as follows

- If all measurements contribute equally to the precision of the location estimate, the improvement in location precision is basically due to a greater number of measurements being available for noise suppression rather than an inherent advantage associated with combinations of different types of measurement.

- From the analysis of the example TRILOC processor, the errors in measuring radial acceleration cause much larger location errors than either measurements of range-rate or range difference.
- The value of range rate and radial acceleration measurements decrease rapidly with increased satellite altitude (e.g. from a two to three hour period) while the value of range difference measurements does not change appreciably.

An overall characteristic of a TRILOC system is the opportunity to improve location precision without sacrificing system capacity or alternatively, to increase system capacity without degrading location precision. If system capacity is dominant, then the acquisition of three measurements of location parameters during a single transmission provides a nearly threefold increase in system capacity compared to current data collection and location systems. On the other hand, if precise location is most important, then the additional location parameter measurements acquired during each transmission can be utilized to improve location accuracy without reducing system capacity of current systems.

The potential advantages and disadvantages of both sweeping and fixed directive antennas on-board the satellite for random access systems are analyzed in Section 3. These analyses indicate that there is no advantage (and probably only disadvantages) in having a sweeping antenna on-board the satellite compared to multiple directive antennas from either a performance or a system capacity viewpoint. This conclusion is based upon a necessary increase in mutual interference between platform transmissions as the gain of a sweeping antenna increases. However, directive antennas do offer advantages in system capacity improvement by decreasing interference, platform power reduction, and/or RFI suppression.

To achieve the advantages of the directive antennas, the beams must be of elliptical cross section and maintained in orientation relative to the satellite's direction of motion. In particular, the major axes of the beams must be parallel to the satellite's sub-track and be able to receive transmissions from platforms from one edge of the satellite's visibility circle to the other. System capacity will be directly proportional to the number of these beams (from an interference standpoint) that are used to completely illiminate the visibility circle-i.e. the minor axes of the beams determine the number of beams required and also the improvement factor.

To determine antenna size, the minor axis of the beams and operating frequency must be specified. To estimate required sizes for a low altitude satellite (Nimbus or TIROS), a synthetic aperature or array is assumed at an operating frequency of 400 MHz. This leads to the approximate relationship that the maximum antenna dimension in meters is about .32 times the number of beams.

2. TRILOC CONCEPT

2.1 Position Determination

The bases for satellite referenced location systems is the accurate knowledge of satellite position and velocity throughout its orbit and the ability to measure range or range rate between the satellite and the platform to be located. For example, in the case of the IRLS system, range is measured at two points during an overpass of the satellite. At each point, this measured range plus known satellite position establishes a sphere centered at the satellite upon which the platform is known to lie. Assuming the platform lies on the earth's surface, then the two range spheres and the earth's surface have two points of simultaneous intersection—one of which corresponds to the geographic coordinates of the platform.

An entirely analogous location technique is that utilized by the RAMS system. Here, the measured quantity between satellite and platform is range-rate which give rise to cones with apexes at the satellite (instead of spheres). The location is accomplished in the same manner, however, by determining the two points of simultaneous intersection of two cones and the earth's surface. A third example is the TRANSIT navigation system that measures the differences in range between satellite and platform to estimate location.

The TRILOC concept is to measure not just range (or range difference), or range-rate but to measure both of these quantities and radial acceleration during any single transmission from a platform. For example, if range and range-rate were measured during a single transmission from a platform, then its location could be estimated by determining the two points of simultaneous intersection of a sphere (range), a cone (range-rate) and the earth's surface (assumed altitude of the platform). Based on this alone then, addition of a range measurement during each transmission of a RAMS platform would effectively halve the number of transmissions required for location purposes.

As mentioned, radial acceleration is also capable of being measured and is as equally useful as range or range-rate for location purposes. In this case, the geometric surface upon which the platform is known to lie can be shown to be shaped like a doughnut with no center hole and whose center is at the satellite's position when the measurement is made.

Conceptually then, as many as three measurements could be obtained from a single platform transmission for purposes of location. Whether this is possible, particularly, in the case of the range measurement, depends upon the communication link between the satellite and platform—i.e., range measurement requires a two way link, as in the IRLS system. On the other hand, measurement of range difference between two transmission points can be accomplished with one way communication links.

2.2 Position and Velocity Determination

The above discussions indicate platform position can be determined with at most half the number of platform transmissions of current location systems by utilizing the TRILOC concept. The more general need, however, (particularly when increased precision of location is required) is the ability to estimate platform velocity as well as location. Furthermore, precise location requires correction for center frequency drift and perhaps drift rate during a satellite overpass. Instead of two measurements being required, then, precise location probably requires four to six measurements during an overpass (if not more to permit noise suppression filtering). Because three measurements are acquired for each transmission received in a TRILOC system, the necessary four to six measurements can be obtained with as few as two transmissions per overpass. Therefore, implementation of the TRILOC concept can result in a factor of three reduction in number of transmissions required per overpass.

2.3 Location/Velocity Estimation

While a geometric interpretation of satellite-referenced location facilitates description of the TRILOC concept, the actual method of deriving location and velocity estimates is accomplished in a different manner. This will be described first by limiting discussion to a system ignoring platform velocity and relying solely on range measurements to estimate location. Secondly, the description will be extended to a TRILOC system ignoring platform velocity. These discussions will then be used as the basis for describing a TRILOC system estimating position and velocity while correcting for long term drift in platform transmitted frequency.

2.3.1 Range Measurement Location

If two measurements of range between the satellite and platform are obtained, the above discussions indicate the location of the platform to be one of two points of simultaneous intersection between two range-spheres and the earth's surface. Another way to establish the platforms location is to compute, a priori, for each possible location of the platform the two corresponding values of range that would have been measured if, in fact, the platform were at that location. When the actual pair of range measurements is acquired, the location of the platform can then be determined by matching the measured pair to a computed pair. Note, except for a usually resolveable ambiguity, only one of the computed pair will match the measured pair.

Actually the necessity to pre-compute all possible pairs of range measurements isn't necessary if the following procedure is followed. First, establish an estimate of the geographic coordinates of the platform (X, Y). This estimate can be as crude as requiring the position to be within view of the satellite when both of the range measurements are made. Secondly, for this estimated position and the position of the satellite when the actual range measurement were made, compute the pair of measured ranges that would have occurred if the estimated location were correct. These will of course differ from the measured ranges—unless the guessed location is correct.

The difference between the measured and computed ranges are then used to correct the X,Y estimate of platform location to more closely correspond to the actual platform location. This is accomplished by equating the difference in each case to the sum of two terms. One of these terms is that portion of the (measured-computed)range difference occurring because the guess of the X coordinate is in error and the other term is the increment caused by the error in the estimated Y coordinate of the platform. Symbolically, this may be written as

$$\left(\frac{\Delta R_1}{\Delta X}\right)\delta x + \left(\frac{\Delta R_1}{\Delta Y}\right)\delta Y = R_{1_{MEASURED}} - R_{1_{COMPUTED}}$$

$$\left(\frac{\Delta R_2}{\Delta X}\right)\delta x + \left(\frac{\Delta R_2}{\Delta Y}\right)\delta Y = R_{2_{MEASURED}} - R_{2_{COMPUTED}}$$

where δx and δy are the errors in the X and Y estimates of platform location

$\left(\frac{\Delta R_1}{\Delta X}\right)$ is the difference in(measured minus computed)range that will occur at the first measurement point if the position estimate were in error by ΔX

$(\Delta R_1/\Delta Y)$, $(\Delta R_2/\Delta X)$ ---- analogous to $(\Delta R_1/\Delta X)$

$\left(\frac{\Delta R_1}{\Delta X}\right)\delta X$ is that portion of$(R_1$ measured - R_1computed$)$caused by a δX error in estimating the X coordinate of platform location.

However, if these two equations are solved simultaneously for δX and δY and these values are used to correct the original X,Y estimate of platform location,then the result will be a better estimate if not the actual location of the platform.

An important extension of this type of solution process concerns those conditions wherein the number of range measurements is greater than two—i.e., the number of measurements is greater than the number of unknowns. If there are no errors present in the range measurements or satellite position, such redundant measurements cannot of course serve any purpose. However,

with errors present, redundant measurements can be utilized to suppress the effects of the errors provided they are random with negligible bias. This is accomplished by forming a sum of the squares of the differences between measured and computed (i.e., based on estimated location) ranges at each measurement point and then determining the location of the platform that minimizes this sum. For the case where the errors are normally distributed, this process can be shown to decrease the error in locating the platform in proportion to the reciprocal of the square root of the number of measurements.

2.3.2 TRILOC Location

Analogous to the above, assume a TRILOC system acquires range and range-rate measurements between the satellite and platform. For the case where two measurements are acquired, the location process described above for a two range-measurement process is equally applicable. However, the extension to the redundant measurement solution is not quite as straight forward.

In the case with redundant measurements of the same type—i.e., all range or all range-rate) minimizing the sum of the squares of measured minus computed difference is conceptually valid. However, for the case where different measurements are involved—e.g., range and range-rate—) minimizing the algebraic sum of the squares of differences is not valid without assigning relative weights to the different measurements, let alone adjusting units to make the sum meaningful.

One means of accomplishing this weighting of the different measurements is to divide each difference by the standard deviation of the error for that measurement. This will, in effect, magnify those differences that correspond to low error measurements relative to those of high error—additionally, the sum of the squares of these quotients becomes the sum of non-dimensional terms—i.e., compatible units. This enhancement of the more precise measurement differences serves to drive the location solution to that indicated by these measurements

Note this also means that when the standard deviation of one type of measurement error is large relative to others, then this measurement is effectively ignored in the solution process. This can be seen in the data presented in Figure 1.

Figure 1 is a presentation of the errors in measuring range difference between successive transmissions from a platform, range rate, and acceleration normalized by the resulting mean location error—i.e., the average magnitude of the location error. A basic assumption utilized to develop these data is that the measurement errors are normally distributed with zero mean. Furthermore, there are no errors present except the measurement errors. These data correspond to a satellite orbit similar to the Nimbus spacecraft. Additionally, the measurements are acquired over two successive overpasses of the satellite wherein the transmissions are separated by five minutes. The geometry of the overpass is such that the sub-tracks are evenly spaced to each side of the platform and under these conditions, a total of 3 transmissions are received on each overpass. Note, this means a total of 3 measurements of range-rate and acceleration are received per overpass, but only 2 measurements of range difference are obtained, for a total of 8 measurements per overpass.

The first point of interest for these data are the asymptotes noted in Figure 1. These boundaries correspond to the condition wherein the contribution of two of the three measurements to the location estimate are ignored because the standard deviations for these measurements are large relative to the third. Note, the asymptote for large errors in range difference and range-rate cannot be shown in Figure 1. For example, if the standard deviations of range-rate and acceleration are large, then the standard deviation of range difference measurements establishes location error —i.e., mean location $= \sigma_{\Delta R}/.56$. Similarly, if range difference and acceleration errors are large, then the location error boundary is $\sigma_{\dot{R}}/3.48 * 10^{-3}$.

A more important result obtainable from Figure 1, however, is the reduction in mean location error when both range difference and range-rate measurements are used—assuming acceleration is not used.

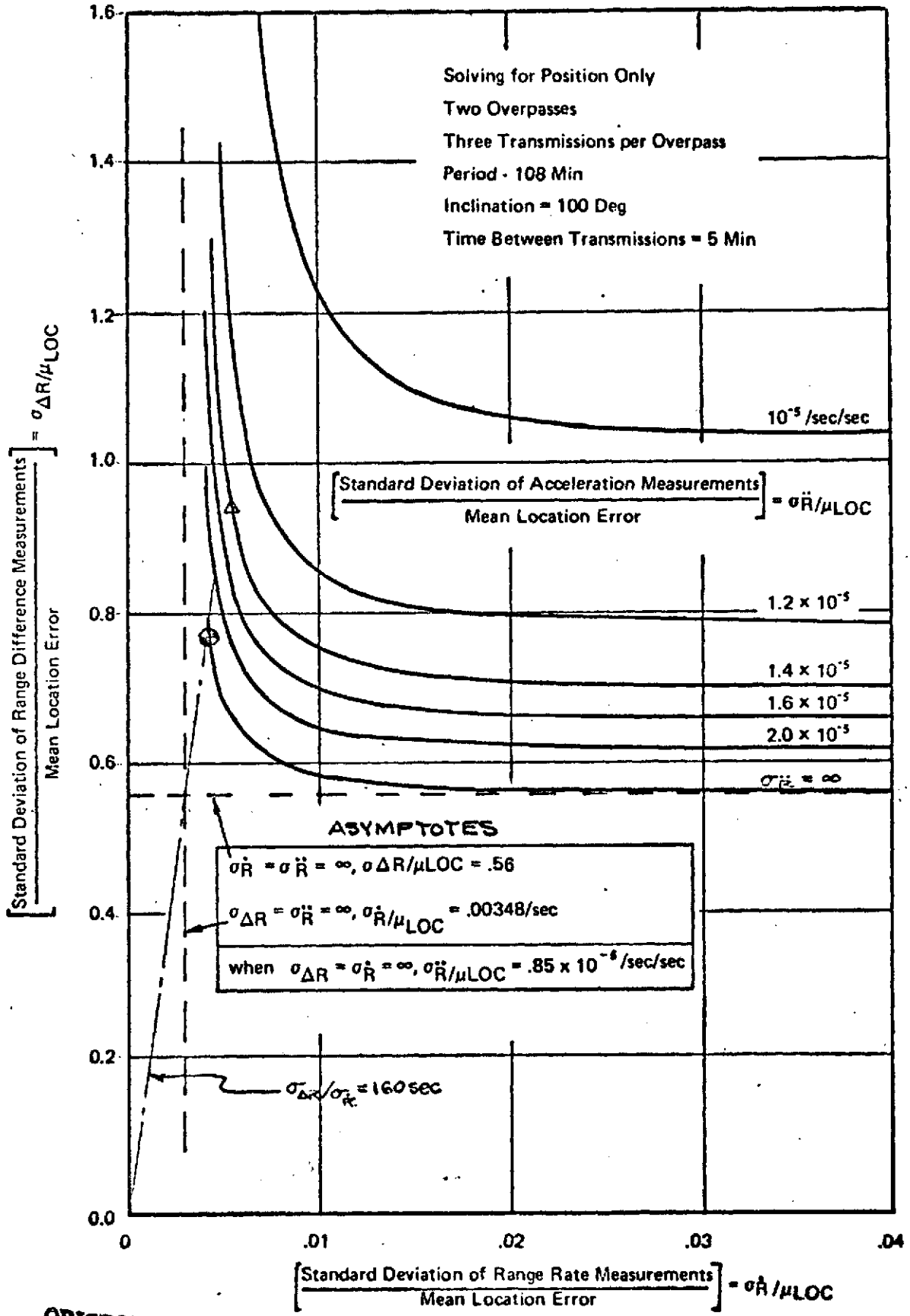

FIGURE 1

To demonstrate the advantage, assume $\sigma_{\Delta R}$ is equal to .56 kilometers and $\sigma_{\dot{R}}$ is equal to 3.48 meters/second. For the asymptotes discussed above, the mean location error, μ_{LOC}, for systems using one or the other of range-rate or range difference measurements is found from:

- Range-rate—the asymptote is described by $\sigma_{\dot{R}}/\mu_{LOC}$ = .00348 per second. With $\sigma_{\dot{R}}$ equal to 3.48 meters/second, μ_{LOC} is then 3.48/.00348 or 1 kilometer.
- Range-difference—the asymptote is described by $\sigma_{\Delta R}/\mu_{LOC}$ = .56. With $\sigma_{\Delta R}$ equal to .56 kilometers, μ_{LOC} is then .56/.56 or 1 kilometer.

To determine the mean location error when both measurements are used simultaneously, manipulation of the data provided in Figure 1 is necessary.

With $\sigma_{\Delta R}$ and $\sigma_{\dot{R}}$ fixed at .56 kilometers and 3.48 meters/second and with $\sigma_{\ddot{R}}$ large (or ignored), the point on Figure 1 that establishes mean location error is the intersection of two lines. One of these lines is the $\sigma_{\ddot{R}} = \infty$ line of Figure 1. The other line is one which must be drawn and is that line along which the ratio between $\sigma_{\Delta R}$ and $\sigma_{\dot{R}}$ is a constant—in this case, $\sigma_{\Delta R}/\sigma_{\dot{R}}$ = .56/.00348 ~ 160 seconds. This line is sketched on Figure 1 and the intersection between this line and the $\sigma_{\ddot{R}} = \infty$ line is noted by the symbol ⊛. At this point, the ordinate and abcissa are:

$$\text{ordinate} - \sigma_{\Delta R}/\mu_{LOC} = .75$$

$$\text{abcissa} - \sigma_{\dot{R}}/\mu_{LOC} = 4.6 \text{ per second}$$

With either the .56 value of $\sigma_{\Delta R}$ or the 3.48 value of $\sigma_{\dot{R}}$, the mean location error is then determined to be about .75 kilometers.

In the asymptotic case using range-rate measurements, a total of six measurements acquired during two successive overpasses are used. By adding the range difference measurements, of which there are four, there is a total of ten measurements being utilized. If the assumption is made that the reciprocal of the square root of the number of measurements determines location error, then the location error would be one kilometer divided by the square root of 10/6 or .77 kilometers. Because this is nearly equal to the .75 kilometer error derived from Figure 1, this indicates that the number of measurements is the major cause of error reduction instead of an inherent advantage associated with different types of measurements. However, it should be noted that if the number of transmissions from a platform are fixed, then the location precision is improved because of the multiple measurements acquired during each transmission.

As another example, the location error resulting when all three measurements are used can be determined in a manner analogous to the above. The assumption is again made that each measurement has a standard deviation of error such that if it alone were used for location, the mean location error would be 1 kilometer. The standard deviation of the acceleration error is therefore as noted in Figure 1— $.85 \times 10^{-6}$/sec. The standard deviations for range difference and range-rate are as before .56 and 3.48×10^{-3}/sec respectively. With these values, the location error is found to be about .6 kilometers. This is established by the intersection of the $\sigma_{\Delta R}/\sigma_{\dot{R}} = 160$ sec line and a curve (not shown) of constant $\sigma_{\ddot{R}}/\sigma_{\dot{R}} = .85 \times 10^{-6}/.00348 = .0025$ sec. This corresponds to the use of 16 measurements—six range-rate, six acceleration, and four range difference. Note again that $1/\sqrt{16/6}$ also gives an estimated error of about .6 kilometers.

2.3.3 TRILOC Performance

Extension of the previous discussions to the case wherein platform velocity and mean frequency of the platform transmitter during an overpass are

also estimated is not significantly different.* Instead of the relationships given in Section 2.3.1 for X, Y determination only, each (measured-computed) difference is now equated to the sum of five terms corresponding to two velocity coordinates (e.g., east and north under the assumption of negligible or known vertical velocity), mean transmitted frequency and again the two position coordinates. Additionally, a parameter held constant in previous performance descriptions is satellite altitude. Because of the potential variations in satellite altitudes, the effects of this parameter should also be evaluated. Therefore, discussions and data provided below are concerned with TRILOC performance in two ways:

- The variation of platform location error with the standard deviations of range difference, range-rate, and acceleration measurements—the units in this case being anticipated units of the actual measurements to be made, namely micro-seconds for range difference, Hertz for range-rate, and Hertz/second for acceleration
- The variation of location error with satellite altitude, holding the number of platform transmissions constant by increasing the interval between transmissions for higher altitude orbits.

To evaluate the effects of satellite altitude on the location errors of a TRILOC system, three altitudes are considered for sun-synchronous orbits—108 minute, 2 hours, and 3 hours. For these orbits, the number of transmissions during which range-difference, range-rate and acceleration are measured

---

* Section 3.4 provides the mathematical derivation in detail, while Section 3.5 analytically describes the means by which location error is computed as a function of measurement errors.

is held fixed at four per overpass with two overpasses. This means there are a total of four acceleration, four range rate and three range difference measurements per pass or a total of 22 measurements over the two successive overpasses. The overpass geometry for these three orbits are sketched in Figures 2, 3, and 4.

In Figure 2, the overpass geometry for the 108 minute orbit is shown. The platform to be located is taken to be at 0° latitude and 0° longitude and, for the computations of the location error, has zero velocity. As noted, the inclination of the orbit is 100°. The successive passes are seen to be symmetrically placed east and west of the platform position and the circles denote the satellite sub-point at the eight transmissions. The transmissions are separated by four minutes. Figures 3 and 4 provide similar information for the geometries of the two and three hour orbits. For the conditions of these figures, little if any geometric dilution of precision should occur in deriving location of the platform.

Figures 5, 6, and 7 present the variation of location error for a TRILOC system operating with the three different orbit altitudes and geometries described in Figures 2, 3, and 4. As previously mentioned, these data correspond to estimation of two position coordinates, two velocity coordinates and the mean transmission frequency of the platform. The presentation is entirely the same as that described in Figure 1 when only platform position coordinates were estimated from the measurements (except for the introduction of the units noted).

As a first point of interpretation, the improved location accuracy of a TRILOC system can be derived from these Figures at the three different altitudes. This is accomplished by assuming the standard deviations of the measurement errors are equal to those for the single-type-measurement system that has the performance of a one kilometer mean location error. For example, in the case of the 108 minute orbit, Figure 5, these standard deviation asymptotes are .89 hertz for frequency (range-rate) measurements, and .0116 Hertz/second for frequency rate (acceleration) measurements. In this case, a TRILOC

Period = 108 Min (Altitude = 1100 Km)

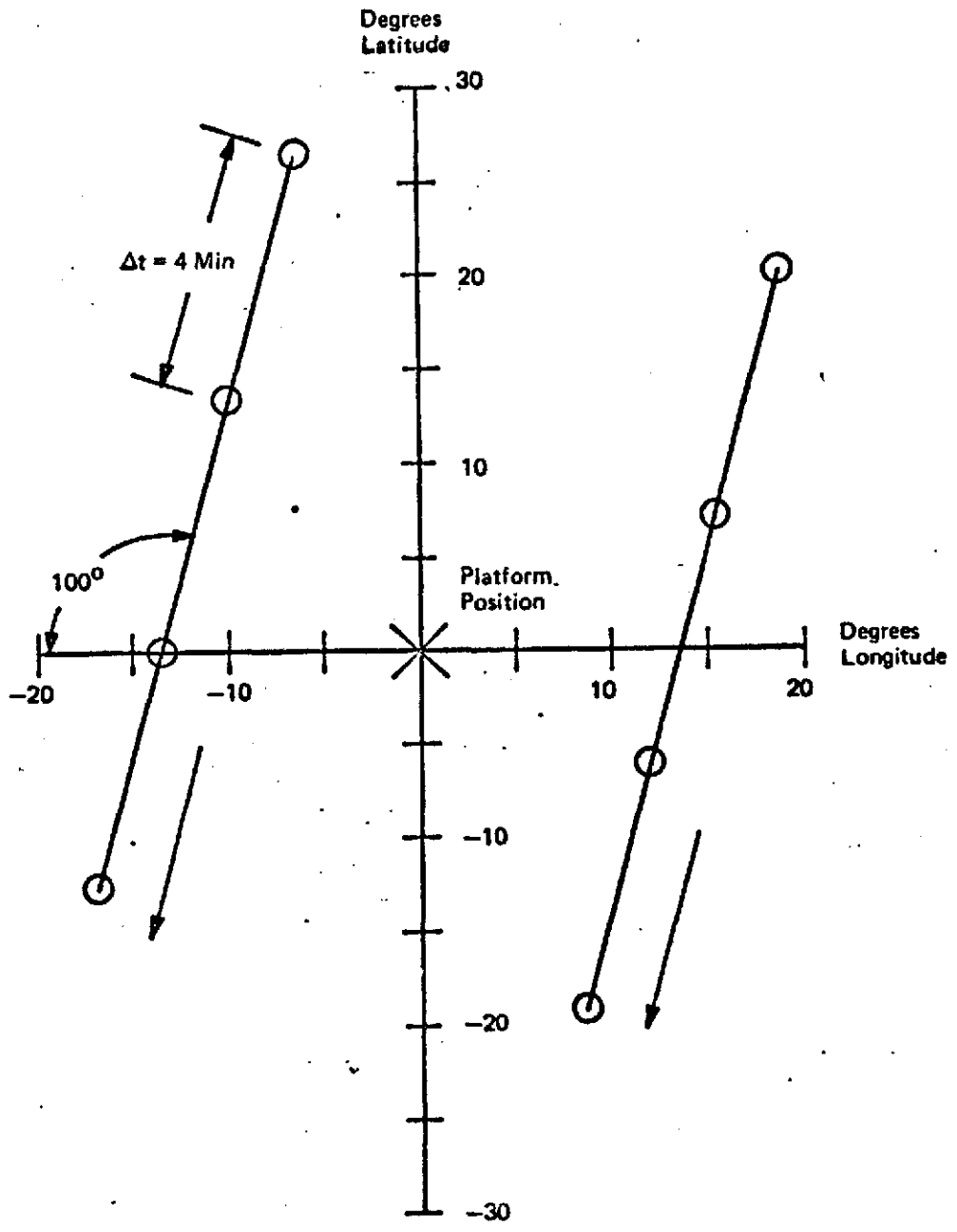

FIGURE 2. OVERPASS GEOMETRY

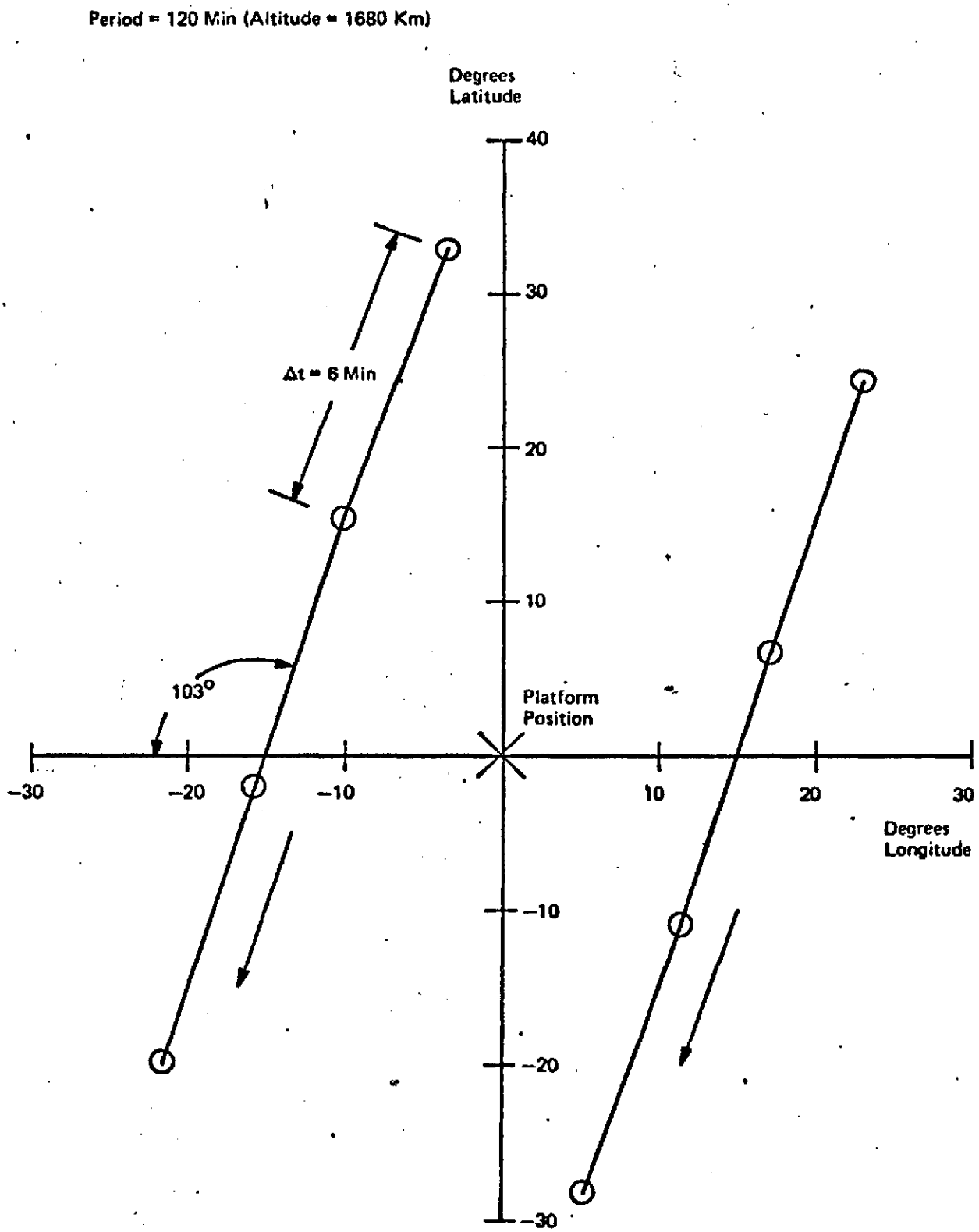

FIGURE 3. OVERPASS GEOMETRY

Pe iod = 180 Min (Altitude = 4182 Km)

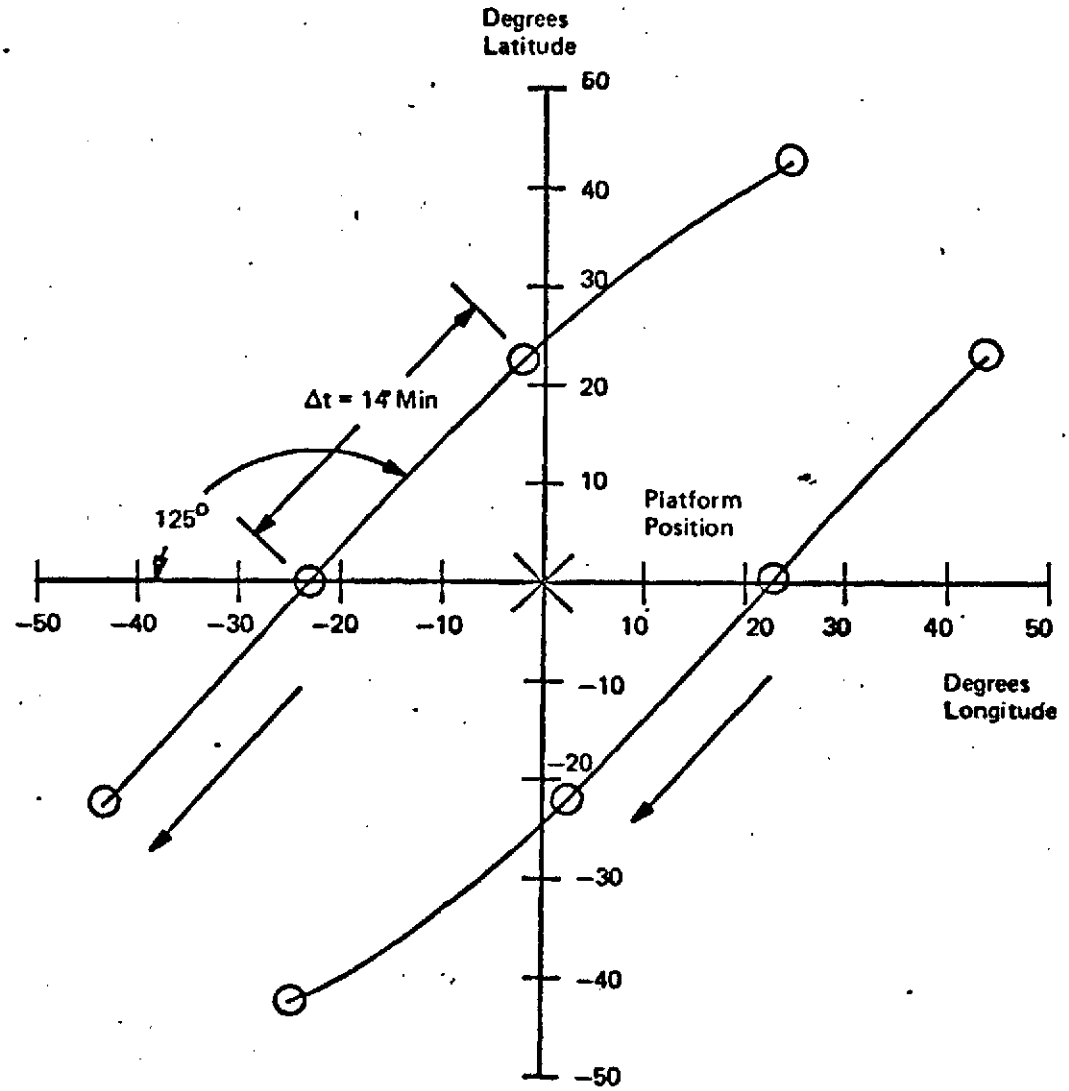

FIGURE 4. OVERPASS GEOMETRY

TRILOC Location Errors

- Solving for
  - Position (Two Coordinates)
  - Velocity (Two Coordinates)
  - Transmission Frequency on Each Overpass
- Overpass Geometry of Figure 2

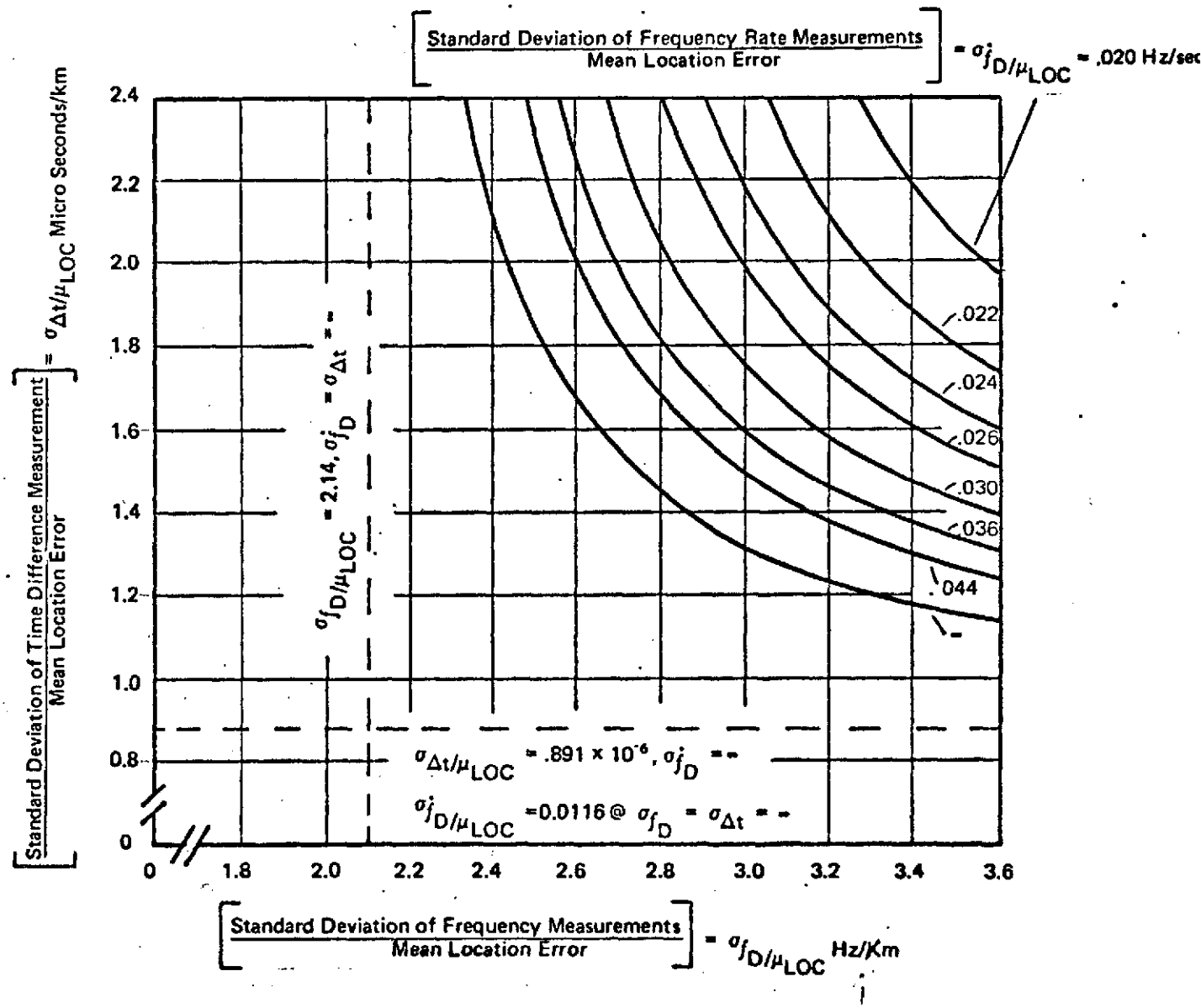

FIGURE 5

TRILOC Location Errors

- Solving for
  - Position (Two Coordinates)
  - Velocity (Two Coordinates)
  - Frequency on each Overpass
- Overpass Geometry of Figure 3

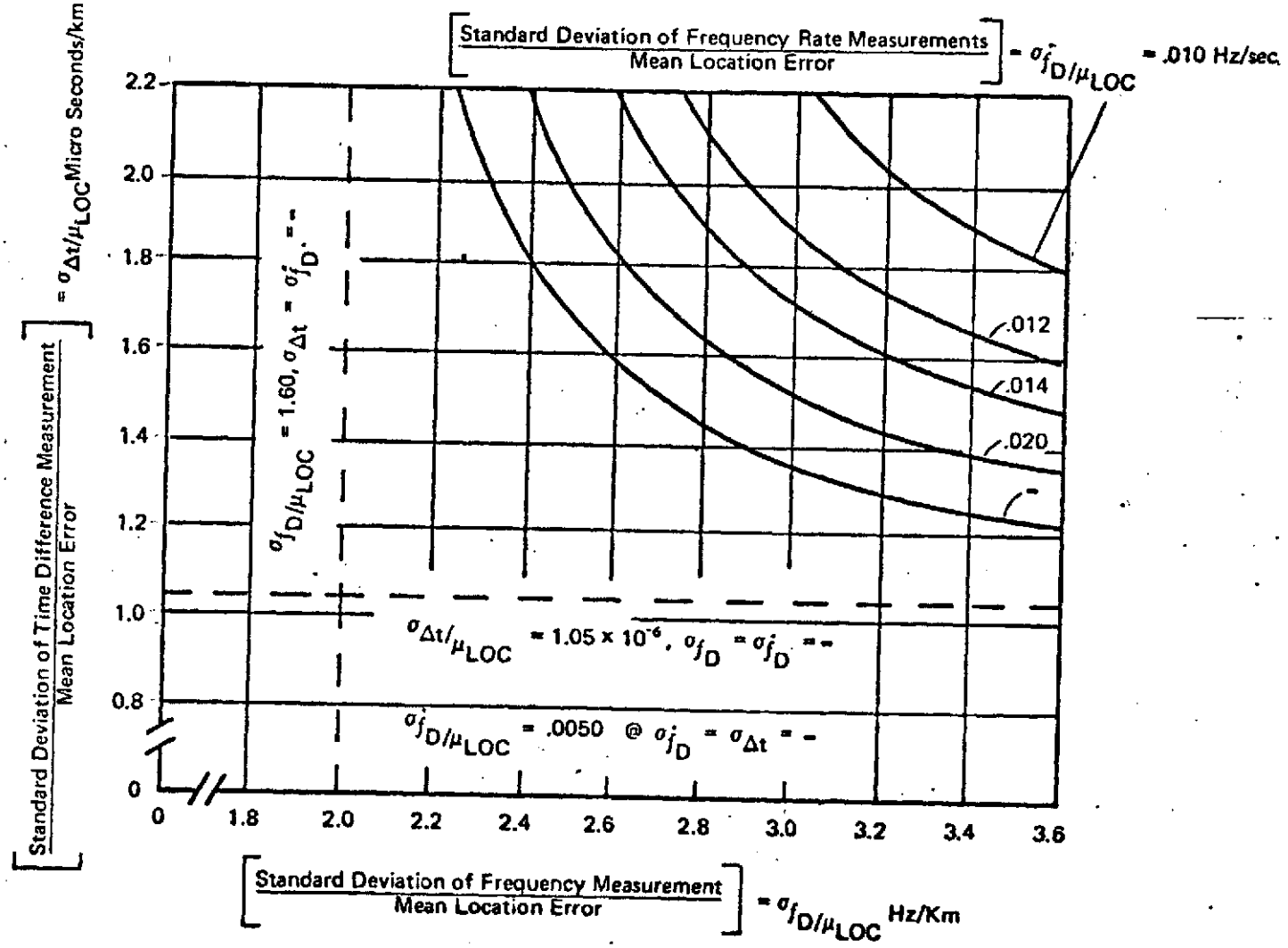

FIGURE 6

TRILOC Location Errors

- Solving for
  - Position (Two Coordinates)
  - Velocity (Two Coordinates)
  - Frequency on each Overpass
- Overpass Geometry on Figure 4

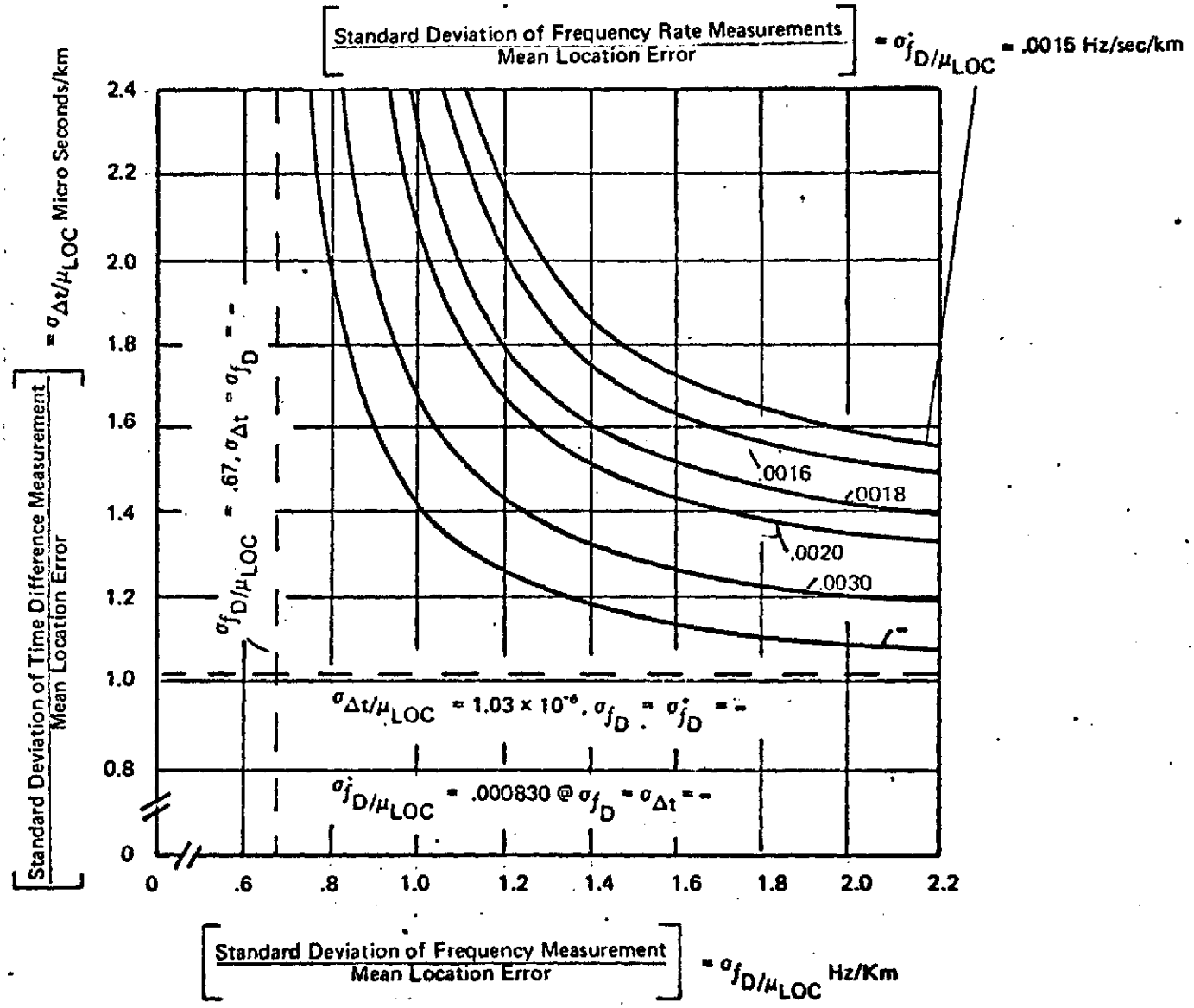

FIGURE 7

system would yield a location error of approximately .53 Km which again is not significantly different than one kilometer divided by $\sqrt{22/8}$.

The variation of location error with altitude may be determined by comparing TRILOC systems with equal measurement errors. For example, if the errors in the above examples are used to enter Figures 6 and 7, then the location errors are found to be .66 Km for the two hour orbit and .81 Km for the three hour orbit. These should then be compared to the .53 Km location error for the 108 minute orbit.

An interesting side light of Figures 5 through 7 concerns the variations of the asymptotes—i.e., the variation of location errors for those systems utilizing only one type of measurement. In the case of both range-rate and acceleration, increasing satellite altitude degrades location performance. For a range rate system, the required standard deviation of error for one Kilometer location error is 2.1 Hertz at 108 minutes and .67 Hertz at 3 hours—or, for given measurement error an increase by a factor of more than three in location error. For an acceleration only system, this factor of increase is about 14. On the other hand, an increase in satellite altitude for a range-difference-only system results in a decrease in location error. This discrepency in location error trends with altitude for the single measurement systems can be qualitatively justified by the following argument.

As the satellite altitude increases, the total variation of both range-rate and acceleration during an overpass will decrease. Furthermore, the variation in acceleration will decrease more rapidly with altitude than range-rate. Therefore, holding measurement errors constant result in relative precision of measurement to the variation of the quantity being measured to decrease—i.e., larger location errors.

This is not true for range difference and in fact, the reverse is true. In this case the total variation of range differences during an overpass increase in going from a 108 minute orbit to a two and three hour orbit. Therefore, the location error for given measurement error should decrease.

A more important aspect of these variations to a TRILOC system is that the higher the satellite altitude, the less advantageous a TRILOC system is. The trends suggest that at some altitude, depending upon measurement errors that are achievable, the location error will be solely determined by the range difference errors—i.e., operation will be limited to the asymptote where the standard deviations of range-rate and acceleration are very large relative to the precision of range difference measurements.

2.3.4 TRILOC System

While the previous discussions indicate the advantages inherent in the multiple measurements of the TRILOC concept, estimates of attainable location precision can only be made by postulating actual mechanisms for measuring the three parameters. To obtain these estimates, a system is suggested in the following discussions regarding the processer on-board a satellite capable of making the requisite measurements. From this, estimates of actual measurement errors are determined and the data of Figure 6 is used to evaluate location precision as a function of system capacity—i.e., number of platforms.

The signal spectrum of this example TRILOC system is a carrier with 20 kHz sub-carriers that are modulated with data. The three parameters necessary for TRILOC location are measured during a transmission interval as follows:

- Received frequency and frequency rate are computed from a least square error regression of measured time intervals between a successive number of positive zero crossings of the carrier while it is tracked by a narrow band phase lock loop

- The range difference measurement is performed by measuring the phase of the 20 kHz sub-carrier relative to a standard on-board the satellite, and comparing this phase to that measured during a successive transmission—phase ambiguities are not a problem in this case because the major portion of any change in range is due strictly to known satellite motion.

To obtain estimates of the errors that can be anticipated in measuring these parameters and the resulting estimate in platform location precision, several assumptions will be made.

- The significant system errors are those associated with the measurements due to noise
- The standards on-board the satellite do not contribute significant error to the measurements
- The overall signal power to noise density ratio is 40 dB-Hz, equally distributed between the carrier and the coherently demodulated subcarrier
- The carrier is tracked with a 20 Hz loop and the subcarrier is measured through a 400 Hz filter
- The measurement period for the carrier is .5 seconds and for the subcarrier is 5 milliseconds.

With these assumptions, the signal to noise ratio for the frequency and frequency rate measurements will be 24 dB while the signal to noise ratio for the subcarrier measurement will be 11 dB. The phase noise in these two channels can then be determined from

For carrier channel: phase noise variance $= \frac{1}{2\ S/N} = .002\ \text{rad}^2$

For subcarrier channel: phase noise variance $= \frac{1}{2\ S/N} = .04\ \text{rad}^2$

Utilizing the regression analysis provided in Appendix F , the standard deviations of the errors in estimating received frequency and its time rate of change are

Standard deviation of frequency error: .0022 Hz

Standard deviation of frequency rate error: .0117 Hz/sec

Where the assumption is made that the nominal frequency is 4 kHz and the period of every cycle is measured and used in the regression analysis to estimate frequency and frequency rate.

Similarly, if the relative phases of 100 positive zero crossings of the 20 kHz subcarrier (5 milliseconds measurement period in order to be able to demodulate the data on the subcarrier during the majority of the transmission period) are measured and utilized to estimate the relative phase, then the standard deviation of the side tone phase measurement will be $\sqrt{.04/100}$ or .02 radians. This translates into

$$\text{Range difference error} = \sqrt{2} \times \text{range error} = \sqrt{2} \times \frac{.02}{2\pi} \times \text{wavelength}$$

For a 20 kHz subcarrier, the wave length is approximately 15 km so that the standard deviations of the range difference error is approximately 70 meters. For purposes of utilizing the data of Figure 6, this is equivalent to .23 microseconds error (i.e., the case where measurement of range difference was based upon signal transit time).

These three errors — .002 Hz, .01 Hz/sec and .23 microseconds— coupled with the data presented in Figure 6 indicates the .002 Hz standard deviation in measuring received frequency will dominate in the location algorithm followed by the range difference measurement and lastly the frequency rate measurement. Therefore, implementation of a TRILOC system versus systems wherein only one parameter is used for location will depend upon the relative values assigned to precise location of platforms and the number of

simultaneous users in a system. Another equally important consideration is the ability to service a mixture of user platforms wherein location precision is important for some and not for others.

For those users requiring precise location, acquisition of a relatively large number of measurements during an overpass is important. If system capacity is limited by interference between platform transmissions, then this increased location precision necessarily results in a reduction in system capacity. This is particularly true for systems wherein only one location parameter is measured during each transmission. However, in a TRILOC system the degradation in system capacity is much slower. Compared to the single measurement system, a threefold increase in number of measurements is available in the TRILOC system for the same number of transmissions. Furthermore, if the example processor previously described is implemented, then the precision with which platforms can be located should be considerably improved compared to current systems typified by IRLS and RAMS.

2.4 TRILOC Location Technique - Analytical

The general problem of locating a platform may be formulated as follows: We want to compute a trajectory parameter vector $\vec{x}$, whose elements may include platform position, velocity, and acceleration, plus parameters relating to transmission frequency. Parameters relating to transmission frequency instabilities may include such items as the center frequency drift and the center frequency drift rate on each satellite overpass. We take a set of m observations, which will be denoted by the column vector $\vec{z}$. In a Doppler location system, for example, the j^{th} component of $\vec{z}$ would be the (observed) received frequency at time t_j. Finally, we assume the existence of an analytical model relating the trajectory parameter vector to the measurement vector, i.e., we assume that there is a known function F such that $F(\vec{x}, t_j) = z_j$, the

j^{th} component of $\vec{z}$. If the dimension (n) of $\vec{x}$ is equal to the dimension (m) of $\vec{z}$, we seek the vector $\vec{x}$ for which

$$F(\vec{x}, t_j) = z_j \tag{1}$$

for j = 1, n.

In practice, however, the dimension of $\vec{z}$ often exceeds the dimension of $\vec{x}$, that is, there are more equations than unknowns. When this occurs, there usually is no vector $\vec{x}$ for which Equation (1) holds for all j because there are always some errors in the observations, and the analytical model represented by the function F may only approximate the physical system. When the number of equations exceeds the number of unknowns, the method of least squares may be used to compute $\vec{x}$.

The least squares technique consists of minimizing the following quadratic form:

$$[\vec{z} - \vec{F}(\vec{x})]^T \Psi^{-1} [\vec{z} - \vec{F}(\vec{x})] \tag{2}$$

In equation (2), $\vec{F}(\vec{x})$ is a column vector with components $F_j = F(\vec{x}, t_j)$, and Ψ is a weighting matrix, which is the covariance matrix of the measurement errors. A necessary condition for the quadratic form (2) to be a minimum is

$$[\vec{D}(\vec{F}(\vec{x}))]^T \Psi^{-1} [\vec{z} - \vec{F}(\vec{x})] = \vec{0} \tag{3}$$

where $\vec{D}$ is the differentiation operator

$$\vec{D} = \left(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_n}\right)$$

The equations (3) are nonlinear and may be transcendental, making an analytical solution difficult or impossible. However, an iterative technique (Newton-Raphson's method) may be used to solve the system of Equations (3). If we have a sufficiently good estimate $\vec{x}_o$ of the trajectory parameter vector, we can expand $\vec{F}(\vec{x})$ into a Taylor series in a neighborhood of $\vec{x}_o$:

$$\vec{F}(\vec{x}) = \vec{F}(\vec{x}_o) + \vec{D}(\vec{F}(\vec{x}_o))(\vec{x} - \vec{x}_o)$$

where we have retained only linear terms.

Equation (3) then become

$$\left[\vec{D}(\vec{F}(\vec{x}_o)\right]^T \psi^{-1}\left[\vec{z} - \vec{F}(\vec{x}_o) - \vec{D}(\vec{F}(\vec{x}_o))(\vec{x}-\vec{x}_o)\right] \Rightarrow \quad (4)$$

To simplify the notation, let $A = \vec{D}(\vec{F}(\vec{x}_o))$. A is then a matrix with elements

$$a_{ij} = \frac{\partial F(\vec{x}_o, t_i)}{\partial x_j}$$

Equation (4) then becomes

$$A^T \psi^{-1}\left[\vec{z} - \vec{F}(\vec{x}_o) - A(\vec{x}-\vec{x}_o)\right] = \vec{0}$$

or

$$\left[A^T \psi^{-1} A\right](\vec{x}-\vec{x}_o) = \left[A^T \psi^{-1}\right](\vec{z} - \vec{F}(\vec{x}_o)) \quad (5)$$

In equation (5), $\vec{x} - \vec{x}_o$ is the error in the estimate of the trajectory parameter vector, i.e., the difference between the initial estimate and the true value of the vector, provided $\vec{x}_o$ is sufficiently close to $\vec{x}$. Similarly, $\vec{z} - \vec{F}(\vec{x}_o)$ is the error in the measurement vector, i.e., the difference between the actual measurement vector ($\vec{z}$) and the value of the measurement vector computed from the initial estimate ($\vec{x}_o$) of the trajectory parameter vector.

The solution may be iterated by beginning with the initial estimate $\vec{x}_o$ and computing successive improved estimates of $\vec{x}$ from

$$\vec{x}_{i+1} = \vec{x}_i + \left[A^T \psi^{-1} A\right]^{-1}\left[A^T \psi^{-1}\right](\vec{z} - \vec{F}(\vec{x}_i))$$

The solution is iterated until successive values of x remain substantially unchanged.

2.4.1 Selection of the Trajectory Parameter Set

If three platform position components and three platform velocity components are to be computed at each time (t_i) of transmission, then each transmission would have to supply at least six independent observables from which position/velocity can be computed. In most systems, only one observable is provided per transmission, and with the TRILOC system at most three are provided. Therefore, a position/velocity estimate at time t_i cannot be obtained

independently of a similar estimate at some different time t_j. Instead, a model of platform motion must be assumed which defines the relationship between platform kinematic parameters at different times. Formally, the model of platform kinematics must have the following form:

$$\vec{R}_B(t) = \vec{R}\left(\vec{R}_B(t_o), \vec{V}_B(t_o), \vec{A}_B(t_o), t - t_o\right)$$

$$\vec{V}_B(t) = \vec{V}\left(\vec{R}_B(t_o), \vec{V}_B(t_o), \vec{A}_B(t_o), t - t_o\right) \qquad (6)$$

$$\vec{A}_B(t) = \vec{A}\left(\vec{R}_B(t_o), \vec{V}_B(t_o), \vec{A}_B(t_o), t - t_o\right)$$

These equations are interpreted to mean that if the kinematic state of the platform is known at time t_o, then its state at any time t may be computed from the functions $\vec{R}$, $\vec{V}$, and $\vec{A}$. In practice, the velocity and acceleration may be assumed to be zero, i.e., the platform is assumed to be stationary. Alternatively, the platform may have a finite velocity but may always remain at a constant (known) altitude. It is this second case that is usually of interest.

A platform at constant altitude may be a stationary platform on the earth's surface, a buoy, or a constant altitude meteorological balloon. If the platform's altitude is constant, then one of its position components is a function of the other two, and it has no vertical velocity component. If the platform follows a great circle route, then its acceleration is a function of its position and its velocity. Thus if the platform moves in a great circle route at constant altitude, the kinematic subset of the trajectory parameter vector contains only two components of position and two components of velocity. Specifically, these components are platform position/velocity components at time t_o. These components are then used to compute the remaining position component, and the entire acceleration vector. Platform position/velocity/acceleration at any other time t is computed from the position/velocity/acceleration at time t_o using Equations (6).

In addition to initial position/velocity estimates, the trajectory parameter vector may also contain components identifying the platform transmission frequency instabilities. Platform location is computed by the information provided by the frequency received at the satellite. Location estimates can be no better than the information on which they are based. Therefore, unless the characteristics of the transmitted signals are known or can be computed to a reasonable degree of precision, the estimates of position and velocity will be inadequate. The instabilities can be computed by assuming an unknown center frequency drift and, possibly, the center frequency drift rate. These unknowns are then computed just as the (unknown) position and velocity are computed. The center frequency drift and drift rate can be computed for each satellite overpass.

2.4.2 The Observables and Their Relation to the Trajectory Parameter Vector

The TRILOC concept includes the measurement three quantities. These are:

1) the received frequency of a signal
2) the time rate of change of the received frequency
3) the difference between the time of reception of two signals (or successive measurements of side tone phase).

These observables are directly related to the relative kinematics between the satellite and the platform.

The first observable is the received frequency, f_R. If the transmitted frequency is f_t, then the Doppler frequency is

$$f_D = f_R - f_t$$

The Doppler frequency is proportional to the time rate of change of the range between the platform and the satellite (V_R):

$$V_R = \frac{dR_R}{dt} = -\frac{c}{f_t} f_D = -\frac{c}{f_t}(f_R - f_t)$$

or,

$$f_R = f_t\left(1 - \frac{V_R}{c}\right) \qquad (7)$$

In Equation (7), c is the speed of light, and f_t is the actual transmitted frequency, which is computed from

$$f_t = f_{t_o} + f_{t_B} + \dot{f}_t\,(t - t_o)$$

where f_{to} is the nominal oscillator center frequency, f_{tB} is the center frequency bias at time t_o, and $\dot{f}_t$ is the center frequency drift rate. V_R is computed from

$$V_R = \frac{d\,|\vec{R}_s - \vec{R}_B|}{dt} = \frac{(\vec{V}_s - \vec{V}_B)\cdot(\vec{R}_s - \vec{R}_B)}{|\vec{R}_s - \vec{R}_B|}$$

where $\vec{R}_s$ and $\vec{V}_s$ are the satellite position and velocity vectors at time t, and $\vec{R}_B$ and $\vec{V}_B$ are the platform position and velocity vectors at time t. $\vec{R}_B$ and $\vec{V}_B$ are computed from the initial estimates of platform position and velocity and Equations (6). Equation (7), therefore, expresses the observable f_R as a function of the trajectory parameter vector, i.e., all or any part of the elements of $\vec{R}_{B_o}$, $\vec{V}_{B_o}$, $\vec{A}_{B_o}$, f_{t_B} and $\dot{f}_t$. Each Doppler measurement produces a relationship of the form of Equation (7).

The second observable is the time rate of change of the received frequency. Its relationship to the trajectory parameter vector is found by taking the time derivative of Equation (7):

$$\dot{f}_R = \dot{f}_t\left(1 - \frac{V_R}{c}\right) - f_t\,\frac{A_R}{c} \qquad (8)$$

Again, $f_t = f_{to} + f_{tB} + \dot{f}_t(t - t_o)$, and c is the speed of light. A_R is the time rate of change of the relative range rate between the satellite and the platform, i.e.,

$$A_R = \frac{dV_R}{dt} = \frac{(\vec{A}_s - \vec{A}_B)\cdot(\vec{R}_s - \vec{R}_B) + |\vec{V}_s - \vec{V}_B|^2 - V_R^2}{|\vec{R}_s - \vec{R}_B|}$$

$\vec{A}_S$ and $\vec{A}_B$ are the acceleration vectors of the satellite and the platform. Equation (8) then expresses the observable $\dot{f}_R$ in terms of the components of the trajectory parameter vector. There is one equation of the form (8) for each TRILOC transmission.

The third observable is the difference in time* between the reception of transmissions(Δt). The time required for a signal to travel from the platform to the satellite is proportional to the relative range (R_R) between the platform and the satellite, where the constant of proportionality is assumed to be the speed of light. If the time (Δt_t) between transmissions from the platform is known, then the difference in relative range (ΔR_R) at two transmissions is proportional to the difference in time between the reception of the two signals, less the difference in time between the transmissions:

$$\Delta R_R = c(\Delta t - \Delta t_t)$$

*The use of successive measurements of side tone phase achieves the same purpose with an entirely analogous derivation.

A possible implementation of this measurement might be a platform that counts zero crossings of the signal provided by its oscillator and is programmed to send a transmission when the number of zero crossings reaches a predetermined number N_c. If the oscillator frequency is f_t then the time between transmission is N_c/f_t. We then have $\Delta R_R = c(\Delta t - N_c/f_t)$ or,

$$\Delta t = \frac{\Delta R_R}{c} + \frac{N_c}{f_t} \qquad (9)$$

ΔR_R is computed from

$$\Delta R_R = |\vec{R}_{S_i} - \vec{R}_{B_i}| - |\vec{R}_{S_{i-1}} - \vec{R}_{B_{i-1}}|$$

where $\vec{R}_{S_i}$ and $\vec{R}_{B_i}$ are the satellite and platform position vectors at time t_i. Equation (9) then relates the observable Δt to the elements of the trajectory parameter vector. The number of equations of the form of Equation (9) on an overpass is one less than the number of transmissions on that overpass. This is because for the first transmission on an overpass there is no previous transmission from which Δt can be computed.

2.5 TRILOC Position Error - Analytical

In Section 2.4 the weighted least squares solution to the TRILOC location problem was derived as

$$\delta\vec{x} = (A^T \psi^{-1} A)^{-1} A^T \psi^{-1} \delta\vec{z} \qquad (1)$$

For convenience, we will write $\vec{x}$ instead of $\delta\vec{x}$ and $\vec{z}$ instead of $\delta\vec{z}$. It is important to remember, however, that $\vec{x}$ and $\vec{z}$ are now the errors in the estimate of the trajectory parameter vector and in the measurement vector, not the vectors themselves. If the measurement errors are sufficiently small, Equation (1) expresses the errors in the trajectory parameter vector as a function of the measurement errors. If ψ is the covariance matrix of the measurement error vector, then

$$\begin{aligned} cov(\vec{x}) &= \left[(A^T\psi^{-1}A)^{-1}A^T\psi^{-1}\right] cov(\vec{z}) \left[(A^T\psi^{-1}A)A^T\psi^{-1}\right]^T \\ &= (A^T\psi^{-1}A)^{-1} \end{aligned}$$

The joint probability density function for the vector $\vec{x}$ is

$$f(\vec{x}) = \frac{|P|^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \exp\left(-\tfrac{1}{2}\vec{x}^T P \vec{x}\right)$$

where $P = (A^T\psi^{-1}A)^{-1}$ and n is the dimension of $\vec{x}$.

We want to find the expected value of the location error, i.e., $E\left\{(x_1^2 + x_2^2)^{\frac{1}{2}}\right\}$ where we have re-arranged the elements of $\vec{x}$ (if necessary) so that x_1 and x_2 are the horizontal components of platform position. Then

$$E\left\{(x_1^2+x_2^2)^{\frac{1}{2}}\right\} = \frac{|P|^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} (x_1^2+x_2^2)^{\frac{1}{2}} \exp\left\{-\tfrac{1}{2}\vec{x}^T P \vec{x}\right\} dx_1 \cdots dx_n$$

The integration is made easier by a transformation to hyperspherical coordinates, (defined in Appendix C):

$$y_1 = \cos\theta_1 \prod_{K=2}^{n-1} \cos\theta_K$$

$$y_i = \sin\theta_{i-1} \prod_{K=i}^{n-1} \cos\theta_K \qquad (2 \le i \le n-1)$$

$$y_n = \sin\theta_{n-1}$$

$$x_i = r\, y_i \qquad (1 \le i \le n)$$

Then

$$\{x_1^2 + x_2^2\}^{\frac{1}{2}} = r \prod_{K=2}^{n-1} \cos\theta_K$$

$$\vec{x}^T P \vec{x} = r^2\, \vec{y}^T P \vec{y}$$

The determinant $|J_n|$ of the transformation from $(r, \theta_1, \ldots, \theta_{n-1})$ coordinates to $(x_1, \ldots, x_n)$ coordinates is (see Appendix C)

$$|J_n| = \left| \frac{\partial(x_1, \ldots, x_n)}{\partial(r, \theta_1, \ldots, \theta_{n-1})} \right| = r^{n-1} \prod_{K=1}^{n-2} \cos^K \theta_{K+1}$$

We then have

$$E\left\{(x_1^2+x_2^2)^{\frac{1}{2}}\right\} = \frac{|P|^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}} \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} \cdots \int_0^{\infty} r^n \left[\prod_{K=2}^{n-1} \cos^K \theta_K\right] e^{-\frac{1}{2} r^2 \vec{y}^T P \vec{y}}\, dr\, d\theta_{n-1} \cdots d\theta_1$$

$$= \frac{|P|^{\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right)}{\sqrt{2}\, \pi^{\frac{n}{2}}} \int_{-\pi}^{\pi} \int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} \frac{\left[\prod_{K=2}^{n-1} \cos\theta_K\right] d\theta_{n-1} \cdots d\theta_1}{\left[\vec{y}^T P \vec{y}\right]^{\frac{n+1}{2}}}$$

Using the results from Appendices D and E, the number of integrations can be further reduced, so that

$$E\left\{(x_1^2+x_2^2)^{\frac{1}{2}}\right\}$$

$$= \left[\frac{|P|^{\frac{1}{2}}\,\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{2}\,\pi^{\frac{n}{2}}}\right]\left[\frac{\pi^{\frac{n+1}{2}}\left\{\prod_{k=3}^{n}\gamma_k^{k-1}\right\}}{2\,\Gamma\left(\frac{n+1}{2}\right)}\right]\int_{-\pi}^{\pi}\frac{d\theta_1}{\left[\vec{y}_2^{\,T}\,P_2\,\vec{y}_2\right]^{3/2}}$$

$$= \frac{|P|^{\frac{1}{2}}\left\{\prod_{k=3}^{n}\gamma_k^{k-1}\right\}}{\sqrt{2\pi}}\int_{0}^{\pi}\frac{d\theta_1}{\left[\vec{y}_2^{\,T}\,P_2\,\vec{y}_2\right]^{3/2}}$$

$$= \frac{|P|^{\frac{1}{2}}\left\{\prod_{k=3}^{n}\gamma_k^{k-1}\right\}}{\sqrt{2\pi}}\int_{0}^{\pi}\frac{d\theta_1}{\left[\alpha_2^2\cos^2\theta_1+2\beta_2\sin\theta_1\cos\theta_1+\gamma_2^2\sin^2\theta_1\right]^{3/2}}$$

The symbols $\alpha_2, \beta_2, \gamma_k$, P_2, and $\vec{y}_2$ are defined in Appendix E.

A closed form solution to the integral was not found, but its value was numerically computed to obtain the results of Figures 5 through 7 previously discussed in Section 2.

3.0 DIRECTIVE ANTENNAS

A directive antenna on-board the satellite may have distinct advantages for data collection and location systems relying on random access. Two such advantages would be the suppression of RFI and/or the ability to reduce radiated power from the platforms. Thirdly, because these systems are limited in capacity by interference between platform transmissions, the use of a directive antenna may serve to alleviate this restriction also.

The analyses that are required to determine the advantages and consequences of the use of a directive antenna(s) can be divided into three categories.

Sweeping Directive Antennas

- Given that a specified number of transmissions from each platform is required during an overpass of the satellite, how are sweep rate and gain of the antenna related to the duty cycle of the platforms?
- Given the relationship between duty cycle and antenna gain, how does increased antenna gain effect the probability of mutual interference between platform transmissions?

Fixed Directive Antennas

- Determine beam shape, gain and number of beams to achieve those advantages apparent from sweeping antennas without the complexity of beam motion.

Antenna Size

- For those antenna gains found to be practical, what are the nominal sizes of antennas that provide the desired gains and beam shapes?

3.1 Sweeping Directive Antennas

3.1.1 Sweep Rate, Gain and Duty Cycle

Referring to the sketch shown in Figure 8, if a satellite is assumed to be at a radius R_s from the center of a spherical earth whose radius is R_E, then the satellite is in line-of-sight view of all platforms within a circle on the earth's surface defined by a cone with apex at the center of the earth, with semi-apex angle of ϵ where $\cos \epsilon = R_E/R_s$.

Having defined this viewing circle, the total solid angle subtended from the satellite to the circumference of this circle defines all directions from which platform transmissions can be received. If this solid angle at the satellite is called A_T, then

$$A_T = 2\pi (1-\sin \epsilon)$$

This solid angle defines a reference point for comparing directive antennas of different gain. In particular, an ideal antenna with a circular beam whose dimensions just cover the visibility circle would have a gain of $4\pi/A_T$. As a numerical example, for a Nimbus satellite, this is an antenna with a gain of about 6 dB.

For a single directive antenna of higher gain than $4\pi/A_T$, complete coverage of the visibility circle will require a sweeping pattern. However, this pattern must be time and position ordered for random access systems. In particular, with platforms transmitting randomly in time from unknown positions within the viewing circle, the motion of the beam cannot be such that by chance, transmissions from a particular platform are missed throughout an overpass. More importantly, where multiple transmissions separated by nominal intervals of time, are necessary, the sweeping pattern becomes further restricted/ordered. The manner in which these requirements can be satisfied is by the imposition of the following:

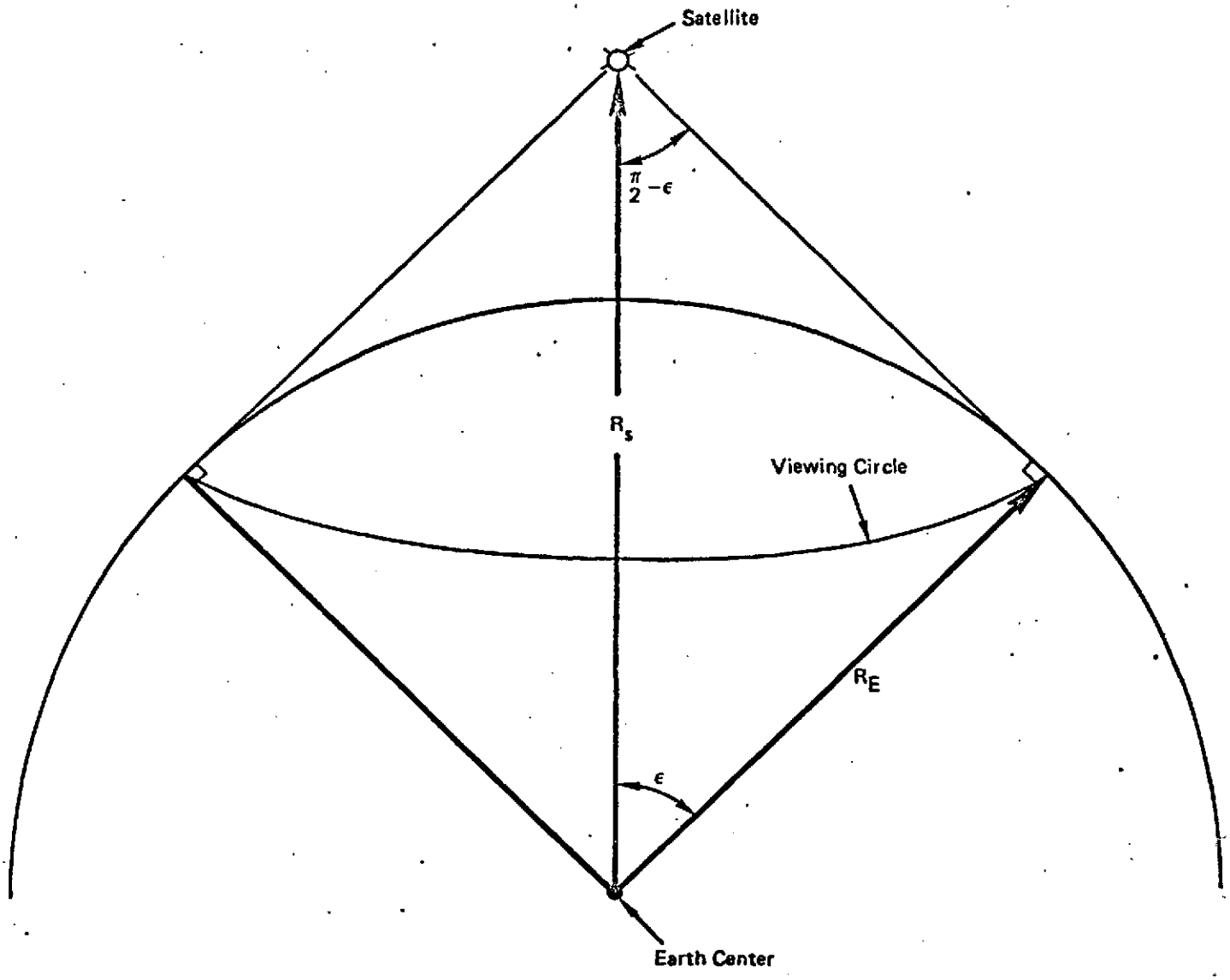

FIGURE 8. SATELLITE VIEWING CIRCLE GEOMETRY

- A directive beam must dwell in any given direction for a period of time greater than the interval between transmissions. This assures receipt of at least one transmission from each platform in the visibility circle during one complete sweep of the visibility circle by the beam.

- The number of complete sweeps of the visibility circle must be no less than the minimum number of transmissions required per overpass from the platforms near the horizon—i.e., those with minimum viewtime.

Satisfaction of these two requirements implies the following. The total viewtime divided by the minimum number of transmissions (plus one) must be equal to the product of the number of discrete positions of the beam and the time interval between transmissions from the platforms. Analytically, these requirements may be written as:

$$\frac{T_V}{(n+1)} = \frac{A_T}{a} T = \frac{G A_T T}{4\pi}$$

where

T_V = period of time the platform is in view of the satellite during an overpass

n = minimum number of transmissions which must be received from the platforms during the viewtime T_V

G = gain of the directive antenna

= $(4\pi/a)$ where "a" is the solid angle subtended by the beam of the directive antenna.

T = Time between transmissions.

In writing these expressions, several simplifying assumptions are made. The ratio (A_T/a) is the number of distinct positions of the beam necessary to completely cover the visibility circle. This means, there is no overlap of coverage from one beam position to another. Furthermore, the shape of the directive beam is such that each beam position is equally effective in covering an area as any other beam position in terms of the projected area of platforms.

For practical antennas, these assumptions are very optimistic in view of the relative geometry let alone the ability to shape beams arbitrarily. However, this assumption does place a lower bound on the number of beam positions and therefore total time required to completely sweep the visibility circle.

By substituting, in the previous expression, the relationship for A_T is namely

$$A_T = 2\pi(1-\sin\epsilon) = 2\pi\left\{1-\sqrt{1-(R_E/R_S)^2}\right\}$$

The time between transmissions can be specified as a function of the gain of the directive antenna beam. That is:

$$T = \frac{2T_V}{(n+1)\,G\left\{1-\sqrt{1-(R_E/R_S)^2}\right\}}$$

As mentioned previously, the minimum gain antenna that can be considered is that one which just covers the visibility circle which is

$$G_{MIN} = \frac{2}{\left\{1-\sqrt{1-(R_E/R_S)^2}\right\}}$$

Therefore, the time between platform transmissions that assures receipt of at least n transmissions during the viewtime T_v may be written as:

$$T = \frac{T_V}{(n+1)} \quad \frac{G_{min}}{G}$$

Another bound on the gain, G, of the directive antenna can be derived from this expression. While G_{MIN} is defined by the antenna that just covers the visibility circle, the maximum antenna gain possible is when the time between transmissions is equal to the duration of transmissions—i.e., the platforms radiate continuously (CW). The possible spread of the antenna gains may then be defined as being between G_{MIN} and the gain value G_{MAX}, corresponding to CW platforms. A numerical example is interesting.

For the Nimbus orbit, the minimum antenna gain is about 6 dB. If RAMS parameters are assumed, —i.e., T_V = 5 minutes, n = 4, and transmission duration of 1 second—then the maximum (CW platforms) antenna gain is about 15 dB. Therefore, the maximum advantage of directive antennas in the RAMS system for purposes of RFI suppression and/or decreased platform power is about 9 dB.

3.1.2 Antenna Gain and Interference

While CW platforms establish an upper bound on antenna gain from the viewpoint of acquiring the requisite number of transmissions, interference between transmissions occurring in the same beam may further limit antenna gain. To evaluate this interference, an assumption regarding distribution of platforms is made. In particular, the platforms are assumed to be uniformly distributed as seen from the satellite.

With this assumption, the number of platforms within the antenna beam of solid angle "a" is Na/A_T—where N is the total number of platforms within line-of-sight of the satellite. The rate at which transmissions from within a beam reach the satellite is then $(Na/A_T T)$, which upon substitution of the above expressions for T, A_T, and "a" leads to:

$$\frac{Na}{TA_T} = N\,\frac{(n+1)}{T_V} = \text{constant}$$

The conclusion, therefore, is that the time rate at which platform transmissions reach the satellite within a given beam is independent of the width of the beam itself. However, this can imply increased interference with higher gain beams as can be seen from the following.

A rough approximation to the distribution of received frequencies at the satellite for platforms uniformly distributed on the earth is that these frequencies are uniformly distributed over a band, F_T. If interference between simultaneously transmitting platforms occurs when they are closer than Δf in frequency, and the transmission arrival times are poisson distributed, then the probability of a given transmission being interfered with can be written as

$$P_I = \text{probability of interference} = 1 - e^{-\frac{2N\tau}{T}\frac{a}{A_T}2\frac{\Delta f}{F_T}}$$

where τ is the duration of platform transmissions. Note, the same assumption is made here, namely the number of platforms that can possibly communicate with the satellite is Na/A_T—i.e., only the fraction (a/A_T) of the total number of platforms as defined by the solid angle "a" of the antenna.

At first glance, this expression would then indicate that the probability of interference would be reduced in direct proportion to the gain of a directive antenna. This is true only if the other parameters remained fixed as antenna gain is varied. However, as shown in Section 3.1, the time between transmissions must be inversly proportional to antenna gain in order to insure receipt of a specified number of transmissions during a given minimum viewtime. By substituting this time and gain relationship, the probability of interference may be rewritten as:

$$P_I = 1 - e^{-\frac{4N(n+1)\tau}{T_v}\frac{\Delta f}{F_T}}$$

Assuming the probability of frequency overlap, $(\Delta f/F_T)$, remains the same, this expression then indicates by the lack of terms describing directive antenna gain, that directional antennas will not influence the probability of interference between platforms.

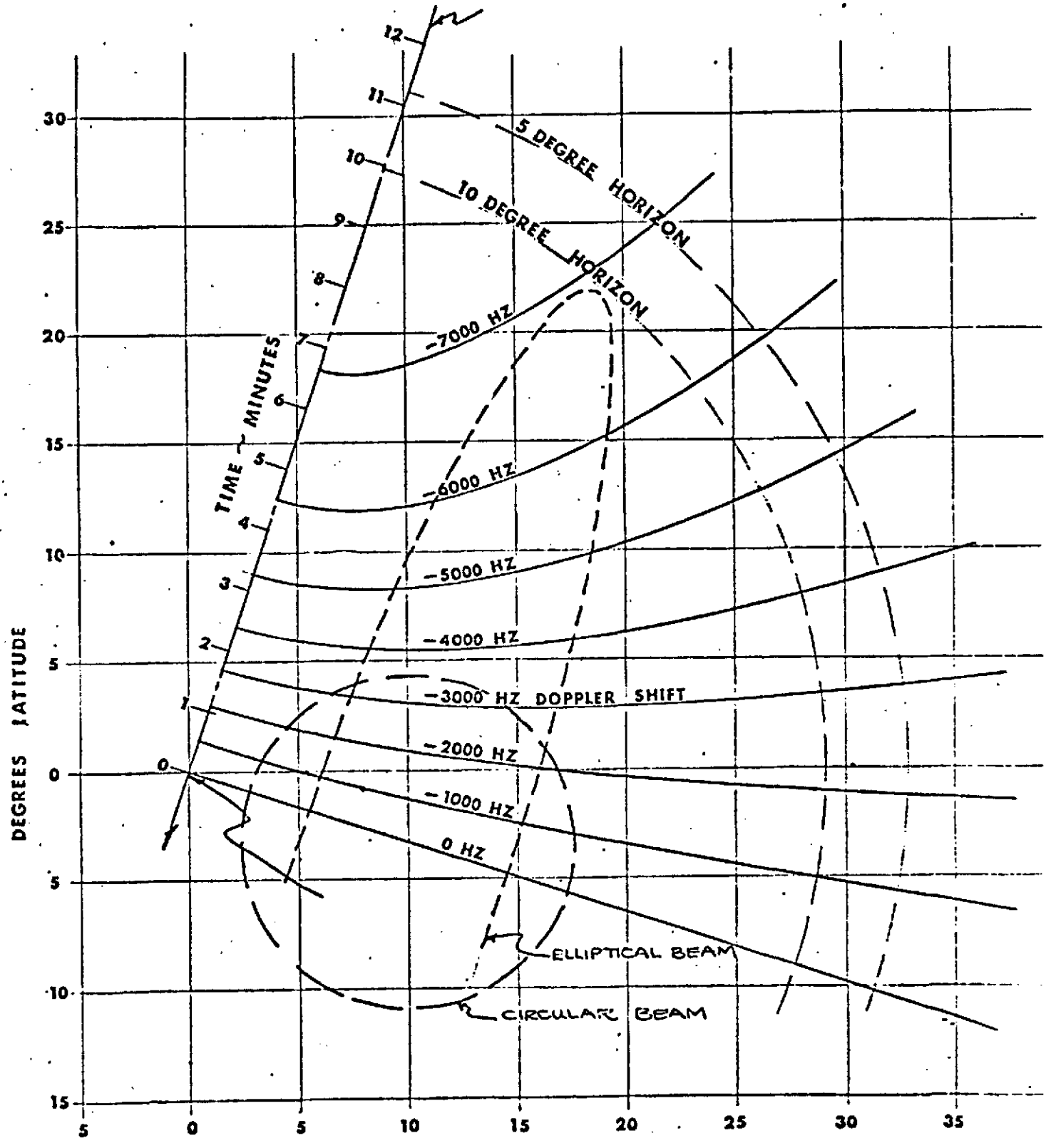

FIGURE 9

900 N.M. SUN SYNCHRONOUS ORBIT

402 MHZ FREQUENCY

Regarding frequency overlap, Figure 9 indicates that the probability of overlap depends upon the shape of the beam of a directive antenna as opposed to the gain. This figure shows, for a platform transmitting at 401 MHz, the locus of platform positions on the earth's surface giving rise to constant doppler shifts in received frequency at the satellite. Also noted are the areas encompassed by two different shaped antenna beams—one circular in shape and the other elliptical with its major axis parallel to the satellite subtrack.

For the elliptical beam, the total range of doppler frequencies can be seen to be about $\pm$ 7000 Hertz or just about the total spectrum of frequencies received if there were a single antenna giving complete coverage of the visibility circle. In this case then, the ratio $\Delta f/F_T$—the probability of frequency overlap—would not be expected to change significantly. Therefore, the probability of interference between platform transmissions would be independent of antenna gain as indicated above.

This conclusion is not true, however, for the circular beam shown in Figure 9. The total spectrum of received frequencies is on the order of $\pm$ 3500 Hertz or about one half the spectrum of frequencies for full coverage. In this case then, $\Delta f/F_T$ would be about twice the value for the elliptic or full coverage beams and the interference between transmissions would be correspondingly higher.

From these considerations, the use of directive antennas implies the requirement to employ beams that are wide parallel to the satellite's subtrack and narrow perpendicular to the subtrack. Otherwise, mutual interference between platforms will increase monotonically with antenna gain.

Because of the rather optimistic assumptions mentioned in the analysis of Section 3.1, it is apparent that sweeping directive antennas cannot basically change mutual interference for random access systems. However, other advantages are possible. In particular, the use of a directive antenna offers at least three opportunities to:

- Maintain platform power and transmission duration constant which, in effect, means utilizing the increased directivity of the antenna strictly as a means to suppress interference from concentrated RFI sources—this might be advantageous even if losses associated with obtaining a directive beam negated any gain increase
- Reduce platform power in inverse proportion to antenna gain in order to simplify the platform
- Reduce the duration of transmissions from platforms in inverse proportion to antenna gain by proportionately increasing platform data rates. This can be shown for random access systems of the RAMS type to simplify signal processing on-board the satellite.

3.2 Fixed Directive Antennas

From the analysis of sweeping antennas, the possibility of reducing interference between platform transmissions is seen to be non-existent because of the necessity to decrease the interval between transmissions in direct proportion to increased gain. Furthermore, because of the optimistic nature of the assumptions regarding coverage of the visibility circle, a sweeping antenna beam would inevitably result in performance degradation in that increased interference could not be avoided as the gain increased. To alleviate this problem, an array of fixed antenna beams might be more advantageous However, the preferred orientation of these fixed beams may be dictated by considerations other than mutual interference.

3.2.1 Beam Shape and Orientation

The orientation and shape of fixed beams will be seen to be governed by two considerations that are conflicting.

- From the viewpoint of satellite signal processing, the antenna patterns should provide coverage such that transmissions are received in a narrow frequency band from each antenna

- From the viewpoint of mutual interference between transmissions, the antenna patterns should provide full coverage of the received frequency spectrum.

In essence, the shape and orientation of fixed beams is a problem of compromise between system capacity in terms of the number of platforms to be serviced and the complexity of satellite processing equipment—basically the complexity of the signal detection function on-board the satellite.

To simplify the detection of transmissions, each fixed antenna should be elliptical in shape with major axis parallel to the zero doppler line—i.e., the major axis of the beam should be perpendicular to the direction of motion of the satellite. The minor axis of the beam's cross section will then determine the band of received frequencies from the platforms within the beam. From the relationship for mutual interference given in Section 3.1.2, this type of beam shape and orientation would not change the probability of interference between platforms compared to an antenna providing full coverage of the visibility circle—i.e., the antenna will accept all those transmissions that can potentially overlap in frequency. However, the detection process is simplified in this case because of the narrower spectrum which needs to be searched for arriving transmissions.

On the other hand, to reduce interference between arriving transmissions the antenna patterns should reject in a geometric sense those platforms which give rise to similar frequencies at the satellite. This can be achieved by orienting elliptic beams with major axis parallel to satellite motion. If the separate beams stretch from one edge of the visibility circle to the other, then the exponent in the relationship for determining probability of interference will

be reduced in near proportion to the number of separate beams. However, the detection mechanism associated with each beam must now cover the full spectrum of received frequencies. For example, if there are equivalently m comb filters required to detect transmissions when a full coverage (single) antenna is used, there will be (mb) comb filters required when there are b separate beams involved.

To analytically evaluate these two orientations of fixed beams, several assumptions analogous to those used to analyze the sweeping antennas will be made. First of all, the band of frequencies received from an antenna is assumed to be proportional to the angular width of the antenna in the direction parallel to the motion of the satellite. Secondly, the number of platforms within a given beam is, as before, assumed to be proportional to the solid angle of the antenna beam.

With these assumptions, the probability of interference may be approximated by

$$P_I = 1 - e^{-\frac{2N\alpha_\perp \alpha_{11} \tau}{T A_T} \frac{2\Delta f}{F_T} \frac{(\pi/2 - \epsilon)}{\alpha_{11}}}$$

where

$\alpha_\perp$ is the angular width of the beam cross-section perpendicular to the motion of the satellite

α_{11} is the angular width of the beam cross-section parallel to the motion of the satellite

so that A_T is the solid angle subtended by the visibility circle of the satellite as before but approximated by a square in this expression whose sides are $(\pi/2 - \epsilon)$.

$\frac{\alpha_\perp \alpha_{||}}{A_T}$ is the fraction of the N platforms within the viewing circle contained within the beam of the antenna

$\alpha_{||}/(\pi/2-\epsilon)$ is the fraction of the total band of frequencies, F_T, received at the satellite that will be contained within the beam of the antenna.

From this expression, the probability of interference is then only a function of the width of the individual beams in a direction perpendicular to the motion of the satellite—i.e., the $\alpha_{||}$ terms cancel. Therefore, for the case where beams have major axes parallel to satellite direction of motion and equal to $(\pi/2-\epsilon)$, the probability of interference may be written as

$$P_I = 1 - e^{-\frac{2N\tau}{bT}\frac{2\Delta f}{F_T}}$$

where b is the number of beams.

3.3 Size of Directive Antenna Arrays

The half power beamwidth of a directive antenna array can be related to size by means of the approximate relationship $\theta = (51\,\lambda/L)$* where θ is the half power beamwidth in degrees, λ is the wavelength of the operating frequency and L is the necessary antenna size perpendicular to the plane within which θ is measured. With this relationship and assuming an operating frequency of 400 MHz (.75 meter wavelength), the maximum antenna dimension is .32b where b is the number of antenna beams of the fixed type recommended in Section 3.2. Note, this last relationship is obtained by assuming a low altitude satellite (Nimbus or Tiros) wherein the total angle which must be covered by the b beams to completely cover the visibility circle is about 120 degrees. As an example then, if there are three equal width antenna beams to each side of the satellite subtrack (i.e., b = 6), then the maximum dimension of the antenna arrays will be nearly two meters.

* Reference-Koelle-Handbook of Astronautical Engineering, page 16-42, McGraw Hill, 1961.

APPENDIX A

PARTIAL DERIVATIVES OF THE MEASUREMENTS

Section 2.4 shows how a platform may be located using an iterative technique for solving certain non-linear equations. Section 2.5 showed how the mean location error may be computed from the measurement errors. Both of these computations involve the use of the partial derivatives of the measurements with respect to the trajectory parameters (e.g., equation (5) of Section 2.4 and equation (1) of Section 2.5). These partial derivates are computed in this Appendix.

A.1 With Respect to Transmission Instabilities

A.1.1 Doppler

In a Doppler location system, the received frequency (f_R) is measured. The relationship between the received frequency f_{R_i} at time t_i and the transmission instabilities is

$$f_{R_i} = f_{t_i} \left(1 - \frac{V_{r_i}}{c} \right) \qquad A.1$$

where f_{t_i} is the frequency transmitted at time t_i, c is the speed of light, and V_{r_i} is the time rate of change of the distance between the platform and the satellite at time t_i. The transmitted frequency, f_{t_i}, is approximated by

$$f_{t_i} = f_{t_o} + f_{t_B} + \dot{f}_t (t_i - t_o)$$

where f_{t_o} is the nominal transmission center frequency, f_{t_B} is the center frequency drift, f_t is the center frequency drift rate, and t_o is a reference time. The reference time t_o may be chosen as the time of the first transmission on an overpass. In the case of a multiple overpass geometry, the frequency drift and drift rate may then be computed for each overpass.

The partial derivatives of f_{R_i} with respect to f_{t_B} and $\dot{f}_t$ are

$$\frac{\partial f_{R_i}}{\partial f_{t_B}} = 1 - \frac{v_{R_i}}{c} \approx 1$$

$$\frac{\partial f_{R_i}}{\partial \dot{f}_t} = (t_i - t_o)\left(1 - \frac{v_{R_i}}{c}\right) \approx t_i - t_o$$

A.1.2 Doppler Rate

In a Doppler rate location system, the time rate of change of the received frequency ($\dot{f}_R$) is measured. The relationship between $\dot{f}_R$ at time t_i and the transmission instabilities is

$$\dot{f}_{R_i} = \frac{d f_{R_i}}{dt} = \dot{f}_t\left(1 - \frac{v_{R_i}}{c}\right) - \frac{f_{t_i} A_{R_i}}{c} \qquad A.2$$

where $A_{R_i} = \frac{dv_{r_i}}{dt}$ is the radial acceleration between the satellite and the platform. Then

$$\frac{\partial \dot{f}_{R_i}}{\partial f_{t_B}} = -\frac{A_{R_i}}{c} \approx 0$$

$$\frac{\partial \dot{f}_{R_i}}{\partial \dot{f}_t} = \left(1 - \frac{v_{R_i}}{c}\right) - (t_i - t_o)\frac{A_{R_i}}{c} \approx 1$$

A.1.3 Range Difference

In a range difference location system, the difference in time (Δt) between received transmissions is measured. The relationship between Δt_i at time t_i and the transmission instabilities is

$$\Delta t_i = \frac{\Delta R_{R_i}}{c} + \frac{N_c}{f_{t_i}} \qquad A.3$$

where ΔR_{R_i} is the difference in range at times t_i and t_{i-1} and N_c is the number of cycles between transmissions. Then

$$\frac{\partial \Delta t_i}{\partial f_{t_B}} = -\frac{N_c}{f_{t_i}^2} \approx -\frac{N_c}{f_{t_o}^2}$$

$$\frac{\partial \Delta t_i}{\partial \dot{f}_t} = -\frac{N_c}{f_{t_i}^2}(t_i - t_o) \approx -\frac{N_c (t_i - t_o)}{f_{t_o}^2}$$

A.2.0 With Respect to Platform Kinematics

Define the following symbols:

1) $\vec{R}_{B_i}$, $\vec{V}_{B_i}$, $\vec{A}_{B_i}$ - platform position, velocity, and acceleration vectors at time t_i

2) $\vec{R}_{S_i}$, $\vec{V}_{S_i}$, $\vec{A}_{S_i}$ - satellite position, velocity and acceleration vectors at time t_i

3) R_{R_i}, V_{R_i}, A_{R_i} = relative range, relative range rate, and relative range acceleration at time t_i.

Platform position, velocity and acceleration at time t_i are functions of the initial state of the platform, i.e., its position/velocity/ acceleration at time t_o:

$$\vec{R}_{B_i} = \vec{R}(\vec{R}_{B_o}, \vec{V}_{B_o}, \vec{A}_{B_o}, t_i - t_o)$$

$$\vec{V}_{B_i} = \vec{V}(\vec{R}_{B_o}, \vec{V}_{B_o}, \vec{A}_{B_o}, t_i - t_o)$$

$$\vec{A}_{B_i} = \vec{A}(\vec{R}_{B_o}, \vec{V}_{B_o}, \vec{A}_{B_o}, t_i - t_o)$$

Each component of all vectors $\vec{R}_{B_i}$, $\vec{V}_{B_i}$, and $\vec{A}_{B_i}$ are functions of each component of the vectors $\vec{R}_{B_o}$, $\vec{V}_{B_o}$, and $\vec{A}_{B_o}$. It is the initial state of the platform that is to be estimated. It will be convenient to use the notation * to denote a component of the initial state of the platform. For example, $\frac{\partial \vec{R}_{B_i}}{\partial *}$

would refer to the partial derivitive of platform position with respect to $R_{B_{x_o}}$ when $* = R_{B_{x_o}}$, with respect to $R_{B_{y_o}}$ when $* = R_{B_{y_o}}$, etc.

A.2.1 Doppler

The relative range rate is

$$V_{R_i} = \frac{(\vec{V}_{s_i} - \vec{V}_{B_i}) \cdot (\vec{R}_{s_i} - \vec{R}_{B_i})}{R_{R_i}}$$

The partial derivatives of V_{R_i} are

$$\frac{\partial V_{R_i}}{\partial *} = -\left[\frac{\frac{\partial \vec{V}_{B_i}}{\partial *} \cdot (\vec{R}_{s_i} - \vec{R}_{B_i}) + (V_{s_i} - V_{B_i}) \cdot \frac{\partial R_{B_i}}{\partial *} + V_{R_i} \frac{\partial R_{R_i}}{\partial *}}{R_{R_i}}\right]$$

where

$$\frac{\partial R_{R_i}}{\partial *} = -\frac{\frac{\partial \vec{R}_{B_i}}{\partial *} \cdot (\vec{R}_{s_i} - \vec{R}_{B_i})}{R_{R_i}}$$

Then

$$\frac{\partial f_R}{\partial *} = -\frac{f_{t_i}}{c} \frac{\partial V_{R_i}}{\partial *} \approx -\frac{f_{t_o}}{c} \frac{\partial V_{R_i}}{\partial *}$$

A.2.2 Doppler Rate

The relative range acceleration is

$$A_{R_i} = \frac{(\vec{A}_{s_i} - \vec{A}_{B_i}) \cdot (\vec{R}_{s_i} - \vec{R}_{B_i}) + |\vec{V}_{s_i} - \vec{V}_{B_i}|^2 - V_{R_i}^2}{R_{R_i}}$$

The partial derivatives of A_{R_i} are

$$\frac{\partial A_{R_i}}{\partial *} = -\frac{1}{R_{R_i}} \left[\frac{\partial \vec{A}_{B_i}}{\partial *} \cdot (\vec{R}_{s_i} - \vec{R}_{B_i}) + (\vec{A}_{s_i} - \vec{A}_{B_i}) \cdot \frac{\partial \vec{R}_{B_i}}{\partial *} \right.$$

$$\left. + 2 \frac{\partial \vec{V}_{B_i}}{\partial *} \cdot (\vec{V}_{s_i} - \vec{V}_{B_i}) + 2 V_{R_i} \frac{\partial V_{R_i}}{\partial *} + A_{R_i} \frac{\partial R_{R_i}}{\partial *} \right]$$

then

$$\frac{\partial \dot{f}_R}{\partial *} = -\frac{\dot{f}_t}{c} \frac{\partial V_{R_i}}{\partial *} - \frac{\dot{f}_{t_i}}{c} \frac{\partial A_{R_i}}{\partial *}$$

$$\approx -\frac{\dot{f}_t}{c} \frac{\partial V_{R_i}}{\partial *} - \frac{\dot{f}_{t_0}}{c} \frac{\partial A_{R_i}}{\partial *}$$

A.2.3 Range Difference

The relative range at time t_i is

$$R_{R_i} = |\vec{R}_{s_i} - \vec{R}_{B_i}|$$

The range difference is then

$$\Delta R_{Ri} = |\vec{R}_{Si} - \vec{R}_{Bi}| - |\vec{R}_{Si-1} - \vec{R}_{Bi-1}| = R_{Ri} - R_{Ri-1}$$

The partial derivatives of R_{R_i} are

$$\frac{\partial R_{Ri}}{\partial x} = - \frac{\frac{\partial \vec{R}_{Bi}}{\partial x} \cdot (\vec{R}_{Si} - \vec{R}_{Bi})}{R_{Ri}}$$

then

$$\frac{\partial \Delta t_i}{\partial x} = \frac{1}{c} \left[\frac{\partial R_{Ri}}{\partial x} - \frac{\partial R_{Ri-1}}{\partial x} \right]$$

APPENDIX B

MODEL OF PLATFORM MOTION

This Appendix presents the model of platform motion used in the TRILOC location error analysis. In this model of platform motion, the platform is at constant altitude with an initial horizontal velocity (which may be zero) and an initial horizontal linear acceleration (which may also be zero). This model approximates a great circle trajectory. The following symbols are defined as:

$\vec{P}, \vec{V}, \vec{A}$ - platform position, velocity, and acceleration at time t.

$\vec{P}_o, \vec{V}_o, \vec{A}_o$ - platform position, velocity, and acceleration at time t_o.

δt - $t - t_o$

C - platform altitude (constant)

Then

$$P_y = P_{yo} + V_{yo}\,\delta t + A_{yo}\,(\delta t)^2$$

$$P_z = P_{zo} + V_{zo}\,\delta t + A_{zo}\,(\delta t)^2$$

$$P_x = \{ C^2 - P_y^2 - P_z^2 \}^{\frac{1}{2}}$$

$$V_y = V_{yo} + A_{yo}\,\delta t \qquad A_y = A_{yo}$$

$$V_z = V_{zo} + A_{zo}\,\delta t \qquad A_z = A_{zo}$$

$$V_x = -\frac{P_y V_y + P_z V_z}{P_x} \qquad A_x = -\frac{V_x^2 + V_y^2 + V_z^2 + P_y A_y + P_z A_z}{P_x}$$

The partial derivatives of $\vec{P}$, $\vec{V}$, and $\vec{A}$ with respect to the components of $\vec{P}_o$, $\vec{V}_o$, and $\vec{A}_o$ are

$$\frac{\partial \vec{P}}{\partial P_{yo}} = \left(-\frac{P_y}{P_x}, 1, 0 \right)$$

$$\frac{\partial \vec{P}}{\partial V_{yo}} = \frac{\partial \vec{P}}{\partial P_{yo}} \delta t$$

$$\frac{\partial \vec{P}}{\partial A_{yo}} = \frac{\partial \vec{P}}{\partial P_{yo}} (\delta t)^2$$

$$\frac{\partial \vec{P}}{\partial P_{zo}} = \left(-\frac{P_z}{P_x}, 0, 1 \right)$$

$$\frac{\partial \vec{P}}{\partial V_{zo}} = \frac{\partial \vec{P}}{\partial P_{zo}} \delta t$$

$$\frac{\partial \vec{P}}{\partial A_{zo}} = \frac{\partial \vec{P}}{\partial P_{zo}} (\delta t)^2$$

$$\frac{\partial \vec{V}}{\partial P_{yo}} = \left(\frac{P_y V_x - P_x V_y}{P_x^2}, 0, 0 \right)$$

$$\frac{\partial \vec{V}}{\partial V_{yo}} = \frac{\partial \vec{P}}{\partial P_{yo}} + \frac{\partial \vec{V}}{\partial P_{yo}} \delta t$$

$$\frac{\partial \vec{V}}{\partial A_{yo}} = \frac{\partial \vec{P}}{\partial P_{yo}} \delta t + \frac{\partial \vec{V}}{\partial P_{yo}} (\delta t)^2$$

$$\frac{\partial \vec{V}}{\partial P_{z_0}} = \left(\frac{P_z V_x - P_x V_z}{P_x^2} \; , 0, 0 \right)$$

$$\frac{\partial \vec{V}}{\partial V_{z_0}} = \frac{\partial \vec{P}}{\partial P_{z_0}} + \frac{\partial \vec{V}}{\partial P_{z_0}} \delta t$$

$$\frac{\partial \vec{V}}{\partial A_{z_0}} = \frac{\partial \vec{P}}{\partial P_{z_0}} \delta t + \frac{\partial \vec{V}}{\partial P_{z_0}} (\delta t)^2$$

$$\frac{\partial \vec{A}}{\partial P_{y_0}} = \left(\frac{P_y A_x - P_x A_y}{P_x^2} - 2 \frac{V_x}{P_x} \left\{ \frac{P_y V_x - P_x V_y}{P_x^2} \right\} , 0, 0 \right)$$

$$\frac{\partial \vec{A}}{\partial V_{y_0}} = \frac{\partial \vec{A}}{\partial P_{y_0}} \delta t - \left(\frac{A_y}{P_x} \delta t - 2 \frac{P_y V_x}{P_x^2} , 0, 0 \right)$$

$$\frac{\partial \vec{A}}{\partial A_{y_0}} = \frac{\partial \vec{P}}{\partial P_{y_0}} + \frac{\partial \vec{A}}{\partial V_{y_0}} \delta t - \left(2 \frac{V_y}{P_x} \delta t, 0, 0 \right)$$

$$\frac{\partial \vec{A}}{\partial P_{z_0}} = \left(\frac{P_z A_x - P_x A_z}{P_x^2} - 2 \frac{V_x}{P_x} \left\{ \frac{P_z V_x - P_x V_z}{P_x^2} \right\} , 0, 0 \right)$$

$$\frac{\partial \vec{A}}{\partial V_{z_0}} = \frac{\partial \vec{A}}{\partial P_{z_0}} \delta t - \left(\frac{A_y}{P_x} \delta t - 2 \frac{P_y V_x}{P_x^2} , 0, 0 \right)$$

$$\frac{\partial A}{\partial A_{z_0}} = \frac{\partial \vec{P}}{\partial P_{z_0}} + \frac{\partial \vec{A}}{\partial V_{z_0}} \delta t - \left(2 \frac{V_z}{P_x} \delta t, 0, 0 \right)$$

The derivatives of $\vec{P}$, $\vec{V}$, and $\vec{A}$ with respect to P_{xo}, V_{xo}, and A_{xo} are zero.

APPENDIX C

n DIMENSIONAL HYPERSPHERICAL COORDINATE TRANSFORMATION

In Section 2.5 the expected location error was written as a multiple integral. It was stated that by performing a transformation from Cartesian coordinates to hyperspherical coordinates the integration could be performed over all variables except one, i.e., the multiple integral could be reduced to a single integral. This Appendix defines the hyperspherical coordinate system and evaluates the determinant of the Jacobian of the transformation from hyperspherical coordinates to Cartesian coordinates.

Define a transformation from n dimensional hyperspherical coordinates $(r, \theta_1, \ldots, \theta_{n-1})$ to n dimensional cartesian coordinates $(x_1, \ldots, x_n)$ as follows:

$$x_1 = r \cos\theta_1 \prod_{k=2}^{n-1} \cos\theta_k$$

$$x_i = r \sin\theta_{i-1} \prod_{k=i}^{n-1} \cos\theta_k \qquad (2 \le i \le n-1)$$

$$x_n = r \sin\theta_{n-1}$$

The inverse transformation is

$$r = \left[\sum_{i=1}^{n} x_i^2 \right]^{\frac{1}{2}}$$

$$\theta_1 = \tan^{-1}\left(\frac{x_2}{x_1}\right)$$

$$\theta_i = \tan^{-1}\left(\frac{x_{i+1}}{\sqrt{x_{i-1}^2 + x_i^2}}\right) \quad (i = 2, n-1)$$

The range of the hyperspherical coordinates is

$$0 \leq r \leq \infty$$

$$-\pi \leq \theta_1 \leq \pi$$

$$-\pi/2 \leq \theta_i \leq \pi/2 \quad (i = 2, n-1)$$

The Jacobian J_n of the transformation is the matrix

$$J_n = \frac{\partial(x_1, \ldots, x_n)}{\partial(r, \theta_1, \ldots, \theta_{n-1})}$$

$$= \begin{bmatrix} \frac{\partial x_1}{\partial r} & \frac{\partial x_1}{\partial \theta_1} & \cdots & \frac{\partial x_1}{\partial \theta_{n-1}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial x_n}{\partial r} & \frac{\partial x_n}{\partial \theta_1} & \cdots & \frac{\partial x_n}{\partial \theta_{n-1}} \end{bmatrix}$$

We want to show that the determant of J_n is

$$|J_n| = r^{n-1} \prod_{i=1}^{n-2} \cos^i \theta_{i+1} \quad \text{for } n \geq 3$$

For n = 3, we have

$$x_1 = r \cos\theta_1 \cos\theta_2$$

$$x_2 = r \sin\theta_1 \cos\theta_2$$

$$x_3 = r \sin\theta_2$$

and

$$|J_3| = \begin{vmatrix} \cos\theta_1 \cos\theta_2 & -r\sin\theta_1 \cos\theta_2 & -r\cos\theta_1 \sin\theta_2 \\ \sin\theta_1 \cos\theta_2 & r\cos\theta_1 \cos\theta_2 & -r\sin\theta_1 \sin\theta_2 \\ \sin\theta_2 & 0 & \cos\theta_2 \end{vmatrix}$$

$$= r^2 \cos\theta_2$$

$$= r^{3-1} \prod_{i=1}^{3-2} \cos^i \theta_{i+1}$$

Therefore, the formula holds for n = 3.

Now suppose

$$|J_{n-1}| = r^{n-2} \prod_{i=1}^{n-3} \cos^i \theta_{i+1}$$

for $n-1 \geq 3$. Define

$$y_1 = r \cos\theta_1 \prod_{K=2}^{n-2} \cos\theta_K$$

$$y_i = r \sin\theta_{i-1} \prod_{K=i}^{n-2} \cos\theta_K \quad (2 \leq i \leq n-2)$$

$$y_{n-1} = r \sin\theta_{n-2}$$

Then $x_i = y_i \cos\theta_{n-1}$ for $i \leq n-1$ and $x_n = r \sin\theta_{n-1}$. The transformation from $(r, \theta_1, \ldots, \theta_{n-2})$ coordinates to $(y_1, \ldots, y_{n-1})$ coordinates is an $(n-1)$ dimensional hyperspherical coordinate transformation. Therefore,

$$\left| \frac{\partial(y_1, \ldots, y_n)}{\partial(r, \theta_1, \ldots, \theta_{n-2})} \right| = |J_{n-1}|$$

$$= r^{n-2} \prod_{i=1}^{n-3} \cos^i \theta_{i+1}$$

If we let $\{a_{ij}\} = J_{n-1}$, i.e.,

$$J_{n-1} = \begin{bmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{bmatrix}$$

then for $i \le n-1$,

$$\frac{\partial x_i}{\partial r} = \frac{\partial (y_i \cos\theta_{n-1})}{\partial r} = \frac{\partial y_i}{\partial r} \cos\theta_{n-1}$$

$$= a_{i1} \cos\theta_{n-1}$$

$$\frac{\partial x_i}{\partial \theta_{j-1}} = \frac{\partial (y_i \cos\theta_{n-1})}{\partial \theta_{j-1}} = \frac{\partial y_i}{\partial \theta_{j-1}} \cos\theta_{n-1}$$

$$= a_{ij} \cos\theta_{n-1} \quad (2 \le j \le n-1)$$

$$\frac{\partial x_i}{\partial \theta_{n-1}} = \frac{\partial (y_i \cos\theta_{n-1})}{\partial \theta_{n-1}} = -y_i \sin\theta_{n-1}$$

$$= -r a_{i1} \sin\theta_{n-1}$$

and

$$\frac{\partial x_n}{\partial r} = \sin\theta_{n-1}$$

$$\frac{\partial x_n}{\partial \theta_{j-1}} = \begin{cases} 0 & \text{, if } j \neq n \\ r\cos\theta_{n-1} & \text{, if } j = n \end{cases}$$

We then have

$$|J_{n-1}| = \begin{vmatrix} a_{11}\cos\theta_{n-1} & \cdots & a_{1,n-1}\cos\theta_{n-1} & -r\,a_{11}\sin\theta_{n-1} \\ a_{21}\cos\theta_{n-1} & \cdots & a_{2,n-1}\cos\theta_{n-1} & -r\,a_{21}\sin\theta_{n-1} \\ \vdots & & & \\ a_{n-1,1}\cos\theta_{n-1} & \cdots & a_{n-1,n-1}\cos\theta_{n-1} & -r\,a_{n-1,1}\sin\theta_{n-1} \\ \sin\theta_{n-1} & \cdots & 0 & r\cos\theta_{n-1} \end{vmatrix}$$

$$= (-1)^{n-2} \begin{vmatrix} a_{11}\cos\theta_{n-1} & -r\,a_{11}\sin\theta_{n-1} & a_{12}\cos\theta_{n-1} & \cdots & a_{1,n-1}\cos\theta_{n-1} \\ a_{21}\cos\theta_{n-1} & -r\,a_{21}\sin\theta_{n-1} & a_{22}\cos\theta_{n-1} & \cdots & a_{2,n-1}\cos\theta_{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n-1,1}\cos\theta_{n-1} & -r\,a_{n-1,1}\sin\theta_{n-1} & a_{n-1,2}\cos\theta_{n-1} & \cdots & a_{n-1,n-1}\cos\theta_{n-1} \\ \sin\theta_{n-1} & r\cos\theta_{n-1} & 0 & \cdots & 0 \end{vmatrix}$$

This determinant is easily evaluated by observing that the cofactors of of the first two elements of the last row are

1) the cofactor C_1 of $\sin\theta_{n-1}$ is

$$C_1 = (-1)^{n+1} \begin{vmatrix} -r\,a_{11}\sin\theta_{n-1} & a_{12}\cos\theta_{n-1} & \cdots & a_{1,n-1}\cos\theta_{n-1} \\ -r\,a_{21}\sin\theta_{n-1} & a_{22}\cos\theta_{n-1} & \cdots & a_{2,n-1}\cos\theta_{n-1} \\ \vdots & & & \\ -r\,a_{n-1,1}\sin\theta_{n-1} & a_{n-1,2}\cos\theta_{n-1} & \cdots & a_{n-1,n-1}\cos\theta_{n-1} \end{vmatrix}$$

$$c_1 = (-1)^{n+1} \left| \begin{bmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{bmatrix} \begin{bmatrix} -r\sin\theta_{n-1} & 0 & \cdots & 0 \\ 0 & \cos\theta_{n-1} & \cdots & 0 \\ 0 & 0 & \cdots & \cos\theta_{n-1} \end{bmatrix} \right|$$

$$= (-1)^{n+1} |J_{n-1}| [-r \sin\theta_{n-1} \cos^{n-2}\theta_{n-1}]$$

2) The cofactor of $r \cos\theta_{n-1}$ is

$$c_2 = (-1)^{n+2} \left| \begin{matrix} a_{11}\cos\theta_{n-1} & a_{12}\cos\theta_{n-1} & \cdots & a_{1,n-1}\cos\theta_{n-1} \\ \vdots & \vdots & & \vdots \\ a_{n-1,1}\cos\theta_{n-1} & a_{n-1,2}\cos\theta_{n-1} & \cdots & a_{n-1,n-1}\cos\theta_{n-1} \end{matrix} \right|$$

$$\left| \begin{bmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{bmatrix} \begin{bmatrix} \cos\theta_{n-1} & 0 & \cdots & 0 \\ 0 & \cos\theta_{n-1} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \cos\theta_{n-1} \end{bmatrix} \right|$$

$$= (-1)^{n+2} |J_{n-1}| [\cos^{n-1}\theta_{n-1}]$$

The determant of J_n is

$$|J_n| = c_1 \sin\theta_{n-1} + c_2 r \cos\theta_{n-1}$$

We then have

$$|J_n| = (-1)^{n-2}\left\{\sin\theta_{n-1}(-1)^{n+1}|J_{n-1}|\left[-r\sin\theta_{n-1}\cos^{n-2}\theta_{n-1}\right.\right.$$

$$\left.\left. + r\cos\theta_{n-1}(-1)^{n+2}|J_{n-1}|\cos^{n-1}\theta_{n-1}\right]\right\}$$

$$= r|J_{n-1}|\left\{\sin^2\theta_{n-1}\cos^{n-2}\theta_{n-1} + \cos^n\theta_{n-1}\right\}$$

$$= r\cos^{n-2}\theta_{n-1}|J_{n-1}|$$

$$= r\cos^{n-2}\theta_{n-1}\left[r^{n-2}\prod_{i=3}^{n-3}\cos^i\theta_{i+1}\right]$$

$$= r^{n-1}\prod_{i=3}^{n-2}\cos^i\theta_{i+1}$$

Thus we have shown by induction that

$$|J_n| = r^{n-1}\prod_{i=3}^{n-2}\cos^i\theta_{i+1}$$

for all $n \leq 3$.

APPENDIX D

Evaluation of $$\int_{-\pi/2}^{\pi/2} \frac{\cos^{p-1}\theta \, d\theta}{\left[\alpha^2\cos^2\theta + 2\beta\sin\theta\cos\theta + \gamma^2\sin^2\theta\right]^{\frac{p+1}{2}}}$$

The mean location error was written as a multiple integral in Section 2.5. In Appendix E the multiple integral is reduced to a single integral. Beginning with θ_{n-1}, the integration over each θ_j leaves an integral of the form

where α, β, and λ are functions only of θ_1, θ_2, and θ_{j-2}. The integral D-1 is evaluated in this Appendix. This result is then used in Appendix E to compute the mean location error.

Note that

$$\int_{-\pi/2}^{\pi/2} \frac{\cos^{p-1}\theta \, d\theta}{\left[\alpha^2\cos^2\theta + 2\beta\sin\theta\cos\theta + \gamma^2\sin^2\theta\right]^{\frac{p+1}{2}}}$$

$$= \int_{0}^{\pi/2} \frac{\cos^{p-1}\theta \, d\theta}{\left[\alpha^2\cos^2\theta + 2\beta\sin\theta\cos\theta + \gamma^2\sin^2\theta\right]^{\frac{p+1}{2}}}$$

$$+ \int_{0}^{\pi/2} \frac{\cos^{p-1}\theta \, d\theta}{\left[\alpha^2\cos^2\theta - 2\beta\sin\theta\cos\theta + \gamma^2\sin^2\theta\right]^{\frac{p+1}{2}}}$$

For convenience, we shall denote the integral to be evaluated by I (θ).

Let $t = \tan\theta$

Then

$$\sin\theta = \frac{t}{[1+t^2]^{\frac{1}{2}}}$$

$$\cos\theta = \frac{1}{[1+t^2]^{\frac{1}{2}}}$$

$$d\theta = \frac{dt}{1+t^2}$$

$t = 0$ when $\theta = 0$ and $t = \infty$ when $\theta = \pi/2$.

Then

$$I(\theta) = \int_0^\infty \frac{dt}{[\alpha^2 + 2\beta t + \gamma^2 t^2]^{\frac{\rho+1}{2}}}$$

$$+ \int_0^\infty \frac{dt}{[\alpha^2 - 2\beta t + \gamma^2 t^2]^{\frac{\rho+1}{2}}}$$

$$= I(t)$$

Now let

$$D = \frac{\alpha^2\gamma^2 - \beta^2}{\gamma^2}$$

$$u = t + \frac{\beta}{\gamma^2}$$

$$v = t - \frac{\beta}{\gamma^2}$$

Then

$$dt = du = dv$$

$$\alpha^2 + 2\rho t + \gamma^2 t^2 = D + \gamma^2 u^2$$

$$\alpha^2 - 2\beta t + \gamma^2 t^2 = D + \gamma^2 v^2$$

$\mu = \frac{R}{\gamma^2}$ and $v = \frac{\beta}{\gamma^2}$ when $t = 0$, and $u = v = \infty$

Then

$$I(t) = \int_{\frac{\beta}{\gamma^2}}^{\infty} \frac{du}{[D + \gamma^2 u^2]^{\frac{\rho+1}{2}}} + \int_{-\frac{\beta}{\gamma^2}}^{\infty} \frac{dv}{[D + \gamma^2 v^2]^{\frac{\rho+1}{2}}}$$

$$= \int_0^{\infty} \frac{du}{[D + \gamma^2 u^2]^{\frac{\rho+1}{2}}} - \int_0^{\frac{\beta}{\gamma^2}} \frac{du}{[D + \gamma^2 u^2]^{\frac{\rho+1}{2}}}$$

$$+ \int_0^{\infty} \frac{dv}{[D + \gamma^2 v^2]^{\frac{\rho+1}{2}}} + \int_{-\frac{\beta}{\gamma^2}}^{0} \frac{dv}{[D + \gamma^2 v^2]^{\frac{\rho+1}{2}}}$$

$$= 2\int_0^{\infty} \frac{du}{[D + \gamma^2 u^2]^{\frac{\rho+1}{2}}} = \frac{2}{D^{\frac{\rho+1}{2}}} \int_0^{\infty} \frac{du}{\left[1 + \frac{\gamma^2 u^2}{D}\right]^{\frac{\rho+1}{2}}}$$

APPENDIX E

Evaluation of $$\int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} \frac{\left[\prod_{k=2}^{n-1} \cos^k \theta_k\right] d\theta_{n-1} \cdots d\theta_1}{\left[\vec{\gamma}^T P \vec{\gamma}\right]^{\frac{n+1}{2}}}$$

In Section 2.5 the mean location error was written as a multiple integral. This Appendix shows how the multiple integral is reduced to a single integral. No analytic expression for this single integral was found but its value can be numerically computed.

$$P = A^T \Psi^{-1} A = \left[cov\left(\vec{X}\right)\right]^{-1}$$

$$\vec{\gamma} = \left[\cos\theta_1 \prod_{k=2}^{n-1} \cos\theta_k, \cdots, \sin\theta_{i-1} \prod_{k=i}^{n-1} \cos\theta_k, \cdots, \sin\theta_{n-1}\right]^T$$

Thus P is an n x n dimensional matrix and $\vec{\gamma}$ is an n dimensional column vector.

Define

$$P_n = P$$

and

$$\vec{\gamma}_n = \vec{\gamma}$$

Then for $j = 2, n$, define

$$P_{j-1} = \gamma_j^2 P_j^* - \vec{r}_j^T \vec{r}_j$$

and

$$\vec{y}_j = \left[\vec{y}_{j-1} \cos\theta_{j-1} \;\vdots\; \sin\theta_{j-1} \right]^T$$

where

$$P_j = \left[\begin{array}{c:c} P_j^* & \vec{r}_j^T \\ \hdashline \vec{r}_j & \gamma_j^2 \end{array}\right]$$

That is, P_j is a j x j dimensional matrix. P_j^* is the (j-1) x (j-1) dimensional matrix derived from P_j by striking out the j^{th} row and the j^{th} column of P_j. $\vec{r}_j$ is a (j-1) dimensional row vector which consists of the j^{th} row of P_j, but without the j^{th} (last) element of that row. γ^2_j is the j^{th} element of the j^{th} row of P_j. $\vec{Y}_j$ is a j dimensional row vector such that

1) $\vec{y}_n = \vec{y}$

2) $\vec{y}_{j-1}$ is computed from $\vec{y}_j$ by striking out the last element of $\vec{y}_j$ and dividing the remaining j-1 elements by $\cos\theta_{j-1}$.

Dashed lines in the expressions above are used to indicate partitioned matrices and partitioned vectors.

Then, if $3 \le j \le n$, we have

$$\vec{y}_j^T P \vec{y}_j = \begin{bmatrix} \vec{y}_{j-1}^T \cos\theta_{j-1} & \sin\theta_{j-1} \end{bmatrix} \begin{bmatrix} P_j^* & \vec{r}_j^T \\ \vec{r}_j & \gamma_j^2 \end{bmatrix} \begin{bmatrix} \vec{y}_{j-1} \cos\theta_{j-1} \\ \sin\theta_{j-1} \end{bmatrix}$$

$$= \left[\vec{y}_{j-1}^T P_j^* \vec{y}_{j-1} \right] \cos^2\theta_{j-1}$$

$$+ 2 \left[\vec{r}_j \vec{y}_{j-1} \right] \sin\theta_{j-1} \cos\theta_{j-1}$$

$$+ \left[\gamma_j^2 \right] \sin^2\theta_{j-1}$$

$$= \alpha_j^2 \cos^2\theta_{j-1} + 2\beta_j \sin\theta_{j-1} \cos\theta_{j-1} + \gamma_j^2 \sin^2\theta_{j-1}$$

where $\alpha_j^2 = \left[\vec{y}_{j-1} P_j^* \vec{y}_{j-1} \right]$

$\beta_j = \vec{r}_j \vec{y}_{j-1}$

Note that α_j^2 and β_j are scalars.

We then have

$$\alpha_j^2 \gamma_j^2 - \rho_j^2$$

$$= \gamma_j^2 \left[\vec{\zeta}_{j-1}^T P_j^* \vec{\zeta}_{j-1} \right] - \left[\vec{r}_j \vec{\zeta}_{j-1} \right]^2$$

$$= \gamma_j^2 \left[\vec{\zeta}_{j-1}^T P_j^* \vec{\zeta}_{j-1} \right] - \left[\vec{\zeta}_{j-1}^T \vec{r}_j^T \vec{r}_j \vec{\zeta}_{j-1} \right]$$

$$= \vec{\zeta}_{j-1}^T \left[\gamma_j^2 P_j^* - \vec{r}_j^T \vec{r}_j \right] \vec{\zeta}_{j-1}$$

$$= \vec{\zeta}_{j-1}^T P_{j-1} \vec{\zeta}_{j-1}$$

We now want to show that

$$\int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} \frac{\left[\prod_{k=2}^{n-1} \cos^k \theta_k \right] d\theta_{n-1} \cdots d\theta_2}{\left[\vec{\zeta}_n^T P_n \vec{\zeta}_n \right]^{\frac{n+1}{2}}}$$

$$= \frac{\pi^{\frac{n-1}{2}} \left[\prod_{k=3}^{n} \gamma_k^{k-1} \right]}{2 \Gamma\left(\frac{n+1}{2}\right) \left[\vec{\zeta}_2 P_2 \vec{\zeta}_2 \right]^{3/2}}$$

Let n = 3. Then (see Appendix D)

$$\int_{-\pi/2}^{\pi/2} \frac{\cos^2\theta_2 \, d\theta_2}{\left[\vec{y}_3^{\,T} P_3 \vec{y}\right]^2}$$

$$= \int_{-\pi/2}^{\pi/2} \frac{\cos^2\theta_2 \, d\theta_2}{\left[\alpha_3^2 \cos^2\theta_2 + 2\beta_3 \sin\theta_2 \cos\theta_2 + \gamma_3^2 \sin^2\theta_2\right]^2}$$

$$= \left[\frac{\sqrt{\pi}\,\Gamma\left(\frac{3}{2}\right)}{\Gamma(2)}\right] \frac{\gamma_3^2}{\left[\alpha_3^2\gamma_3^2 - \beta_3^2\right]^{3/2}}$$

$$= \frac{\left[\pi\right]^{\frac{1}{2}}\left[\frac{\pi}{2}^{\frac{1}{2}}\right]}{\Gamma\left(\frac{3+1}{2}\right)} \frac{\gamma_3^2}{\left[\vec{y}_2^{\,T} P_2 \vec{y}_2\right]^{3/2}}$$

$$= \frac{\pi^{\frac{3-1}{2}}\left[\prod_{k=3}^{3} \gamma_k^{k-1}\right]}{2\Gamma\left(\frac{3+1}{2}\right)\left[\vec{y}_2^{\,T} P_2 \vec{y}\right]^{3/2}}$$

Therefore, the formula holds for n = 3. Now suppose it holds for $n - 1 \geq 3$. Then

$$\int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} \frac{\left[\prod_{\kappa=2}^{n-1} \cos^{\kappa}\theta_{\kappa}\right] d\theta_{n-1} \cdots d\theta_2}{\left[\vec{y}_n^{\,T} P_n \vec{y}_n\right]^{\frac{n+1}{2}}}$$

$$= \int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} \frac{\left[\prod_{\kappa=2}^{n-2} \cos^{\kappa}\theta_{\kappa}\right] \cos^{n-1}\theta_{n-1}\, d\theta_{n-1} \cdots d\theta_2}{\left[\alpha_n^2 \cos^2\theta_{n-1} + 2\beta_n \sin\theta_{n-1}\cos\theta_{n-1} + \gamma_n^2 \sin^2\theta_{n-1}\right]^{\frac{n+1}{2}}}$$

$$= \int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} \frac{\left[\sqrt{\pi}\,\Gamma\left(\frac{n}{2}\right)\gamma_n^{n-1}\right]\left[\prod_{\kappa=2}^{n-2} \cos^{\kappa}\theta_{\kappa}\right] d\theta_{n-2} \cdots d\theta_2}{\Gamma\left(\frac{n+1}{2}\right)\left[\alpha_n^2\gamma_n^2 - \beta_n^2\right]^{\frac{n}{2}}}$$

$$= \frac{\sqrt{\pi}\,\Gamma\left(\frac{n}{2}\right)\gamma_n^{n-1}}{\Gamma\left(\frac{n+1}{2}\right)} \int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} \frac{\left[\prod_{\kappa=2}^{n-2} \cos^{\kappa}\theta_{\kappa}\right] d\theta_{n-2} \cdots d\theta_2}{\left[\vec{y}_{n-1}^{\,T} P_{n-1} \vec{y}_{n-1}\right]^{\frac{n}{2}}}$$

$$= \left[\frac{\sqrt{\pi}\,\Gamma\left(\frac{n}{2}\right)\gamma_n^{n-1}}{\Gamma\left(\frac{n+1}{2}\right)}\right]\left[\frac{\pi^{\frac{n-2}{2}}\left[\prod_{\kappa=3}^{n-1} \gamma_{\kappa}^{\kappa-1}\right]}{2\,\Gamma\left(\frac{n}{2}\right)\left[\vec{y}_2^{\,T} P_2 \vec{y}_2\right]^{3/2}}\right]$$

$$= \frac{\pi^{\frac{n-1}{2}}\left[\prod_{\kappa=3}^{n} \gamma_{\kappa}^{\kappa-1}\right]}{2\,\Gamma\left(\frac{n+1}{2}\right)\left[\vec{y}_2^{\,T} P_2 \vec{y}_2\right]^{3/2}}$$

Thus we have shown by induction that the formula holds for all $n \geq 3$.

APPENDIX F

MEASURING FREQUENCY AND FREQUENCY RATE

This appendix describes one technique whereby the frequency received by a satellite and its time rate of change can be estimated and determines the precision to which these two parameters can be measured. The basic assumption throughout the analysis is that white noise is the only error source corrupting the measurements.

The measurements from which frequency and frequency rate are estimated are the series of elapsed times between positive going zero crossings of the received frequency (after down converting the r.f. signal to several thousands of Hertz). For present purposes, the resolution and precision of these time interval measurements is assumed to be much smaller than the equivalent phase jitter of the signal itself as caused by noise.

Assuming the received frequency is varying linearly with time, then the period of time (T_i) between i positive zero crossings can be related to frequency (w) at a given point in time and the time rate of change of frequency (α) by

$$wT_i + .5\alpha T_i^2 = 2\pi i$$

If perturbations are taken from this equation, then the errors in α, w, and Ti ($\delta\alpha$, δw, δTi) are related by

$$T_i\,\delta w + .5\,T_i^2\,\delta\alpha = -(w+\alpha T_i)\,\delta T_i$$

or for an ensemble of measurements of T, this may be written in matrix notation as

$$[A]\{\delta x\} = \{T\}$$

where for the ith measurement

$$a_{i1} = -T_i/(w+\alpha T_i)$$

$$a_{i2} = -.5\,T_i^2/(w+\alpha T_i)$$

$$\{\delta x\} = \begin{Bmatrix}\delta w\\ \delta\alpha\end{Bmatrix}$$

$$\{\delta T\} = \begin{bmatrix}\delta T_1\\ \vdots\\ \delta T_k\end{bmatrix}$$

With k>2, the errors in δw and $\delta\alpha$ can be determined if the solution for w and α is assumed to be based upon least square fitting or regression. In this case, the solution is

$$\delta x = (A^TA)^{-1}A^T\delta T$$

By assuming the δT's are independent (which they are not strictly speaking) of zero mean and equal standard deviation of σ_T, then the covariance matrix of δx becomes

$$\mathrm{cov}(\delta w, \delta\alpha) = \sigma_T^2\,(A^TA)^{-1}$$

This relationship can be used then to evaluate the standard deviations of w and x (namely σ_w and σ_α) as a function of w and α. This has been done by means of a computer program with the following results.

With frequencies (w) and frequency rates (α) between 1 to 10 kHz and 2 to 100 Hz/sec, the standard deviations of frequency and frequency rate are essentially independent of frequency rate and approximately proportional to reciprocal of the square root of the frequency itself when measurement duration is fixed. For a duration of .5 seconds,

$$\sigma_w \sqrt{S/N} = .07/\sqrt{w} \text{ Hz}$$

$$\sigma_\alpha \sqrt{S/N} = .37/\sqrt{w} \text{ Hz/sec}$$

where f is the nominal or average frequency in units of kilohertz and σ_T is replaced by the relationship between signal to noise ratio and frequency —

$$\sigma_T = \frac{1}{w\sqrt{2S/N}} \qquad \text{for w in radians/sec}$$